Evaluation of Florida's Corporate Tax Credit Scholarship Program
Baseline Report – Compliance and Test Scores in 2006-07

David N. Figlio
Department of Economics, University of Florida
and
National Bureau of Economic Research

Revision: March 6, 2008

I. Background

As required by the Florida Statutes, s. 220.187(9)(j), the Florida Department of Education selected the University of Florida as the Independent Research Group and Professor David Figlio as the Project Director to evaluate Florida's Corporate Tax Credit Scholarship Program. The evaluation plan calls for a series of reports over two years that will, once appropriate data are collected after the 2007-08 academic year, compare student learning gains for program participants to otherwise similar non-participants; study differential family satisfaction between program participants and non-participants; describe the information that individuals used when deciding whether to participate in the program, and if they chose to participate, which schools to select; and measure the degree to which the Corporate Tax Credit Scholarship Program affects public school performance. The 2007-08 school year is the first school year in which it is possible to fully control the collection of test scores of program participants, and is also the first year in which it is possible to measure student test score gains in the participating private schools, a measure required by the legislation. Therefore, the analysis of the effects of the program on recipients and the public schools in general will occur following the administration and collection of 2007-08 school year test scores.

This first report provides baseline data on test scores and private school compliance in testing students who participated in Florida's Corporate Tax Credit Scholarship Program. This baseline report has two principal purposes. The first main purpose is to establish the 2006-07 school year as an appropriate baseline for analyzing the effects of the program. Since the tests analyzed in the baseline report were

administered in the school year prior to the beginning of the evaluation, it is especially important to ensure that the testing complied with state law and that the test scores could be recovered and converted into an analyzable format. Were this to not be possible, it would then be necessary to establish the 2007-08 school year as the baseline year for the evaluation. However, if the data were sufficiently complete for evaluation, it would be possible to use 2006-07 as the baseline year, thereby permitting the reporting of estimated program effects as early as the end of 2008.

The second main purpose of this baseline report is to present basic comparative data regarding CTC Scholarship program participants relative to non-participating students, including test scores and demographic information. While it will only be possible to measure the effects of program participation following the 2007-08 school year, these baseline comparisons can still provide very valuable information.

II. Establishing the baseline for analysis

Compliance

As required by s. 330.287(8)(c)(2), participating schools administered to students an approved nationally norm-referenced test as identified by the Florida Department of Education, including the Stanford Achievement Test, Basic Achievement Skills Inventory, Metropolitan Achievement Test, Iowa Test of Basic Skills, Terra Nova, or the Preliminary Scholastic Aptitude Test and ACT/PLAN (for students in high school

grades) or made provisions for participating students to take statewide assessments at a public school in accordance with s. 220.187(7)(e).

Pursuant to s. 220.187(8)(c)(2), in fall 2007 the Independent Research Organization contacted the 833 private schools that had participating students in grades three through ten during the 2006-07 school year. The Florida Department of Education provided the Project Director with a list of all participating students in 2006-07; of these, 9,721 were in the relevant grades, according to the state records. Schools were provided lists of the relevant students and were instructed to submit test scores to the Independent Research Organization.

In over 99 percent of cases, schools submitted photocopies of official score sheets provided to them by the relevant testing company (e.g., Harcourt). In a small number of schools, the schools scored the tests themselves and forwarded to the Project Director detailed information regarding the nature of test administration and scoring. Schools were requested to provide explanations for any students who had attended school in the relevant grades but for whom they did not submit a valid test score. Because the test scores were for last academic year, a large number of students have since changed schools or left the program. In these cases, the Project Director and his staff made every attempt to acquire these test scores from the student's current school, in the event in which the test-administering school had forwarded the records to the new school. Test scores were double-entered, the original score sheets destroyed, and the resulting electronic databases stored in accordance with s. 1002.22(3)(d)(5) of the Florida Statutes.

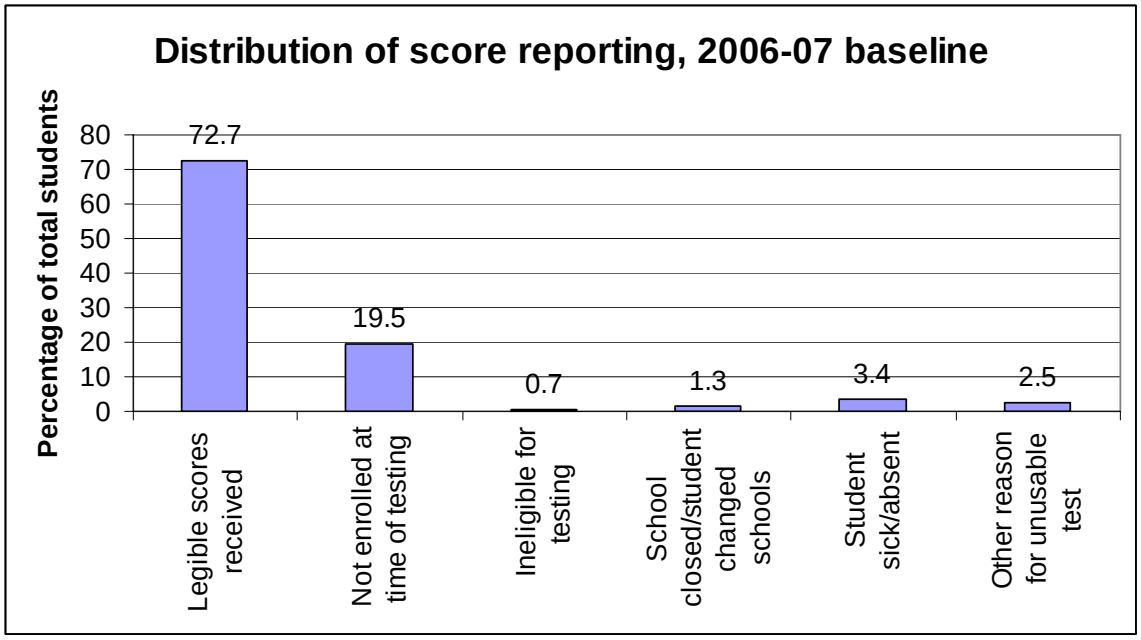
These data were then matched with student FCAT, public schooling, subsidized lunch and disability history, when available, from the Education Data Warehouse, and with information from student scholarship applications provided by the Scholarship Funding Organizations, and de-identified for the purposes of analysis.

Of the 833 schools with students in the relevant grades in 2006-07, 18 schools either closed or left the program; in some of these cases the Project Director was still able to retrieve some of last year's test scores from the students' current schools. This left 815 schools that had students in the relevant grades last year and continue to participate in the program. At the time of writing, every one of these schools provided evidence of test administration. In the case of six schools, problems with test reporting made it impossible to compare students to national norms.

Of the 9,721 students in relevant grades participating in the program in 2006-07, the Independent Research Organization received valid, legible test scores for 7,067 students. In these cases, 6,692 students took only the school's standardized test, 37 took both the statewide assessment and the school's test, and 338 took the statewide assessments.[1] The vast majority of the remainder were not enrolled in the program at the time of testing; 19.5 percent of the 9,721 students potentially eligible for testing either left the program prior to test administration or arrived in the school following test administration (or in a handful of cases, transferred from a late-testing school to an early-testing school midyear), and another 0.7 percent of students were found to be ineligible

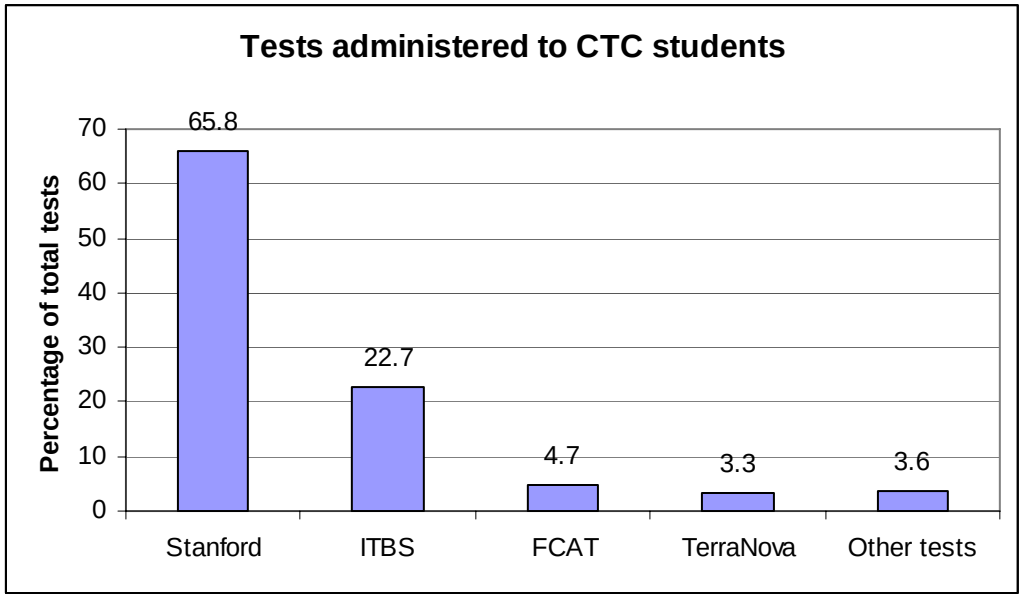
[1] Some of the FCAT-takers took the FCAT while still enrolled in the private school, while others may have returned to a public school by the time of FCAT administration.

for testing pursuant to s. 330.287(8)(c)(2). This left 7,764 testing-eligible students enrolled in the program at the time of testing; these students are potentially eligible to take tests under the program.



Among the remaining students, the least information is known about the 1.3 percent of students for whom the 2006-07 school closed or the student changed schools, but the new school did not receive 2006-07 testing records from the old school. Most of the remaining students (3.4 percent of students) were sick or absent at the time that their school administered the test; in all of these cases, the school demonstrated that they did administer tests to other participating students at a designated time. In the remaining 2.5 percent of cases, there were problems with the test scores reported; either the test scores were reported but illegible or the school provided an incomplete test reporting. More detail on test administration and score reporting can be found in Table II.1. This leaves 7,067 as the number of legible test scores received by the Independent Research Organization.

Table II.2 reports the distribution of tests taken by participating students. As can be seen in the table, 99.54 percent of the students with test scores observed took one of the tests identified by the Florida Department of Education as acceptable for the program. The remaining 33 students took a smattering of other tests, but typically the Woodcock Johnson Psychoeducational Battery. Just over 70 percent of the students took either the FCAT or the Stanford Achievement Test, while another 22.71 percent took the Iowa Test of Basic Skills. The only other test that was taken by more than about one percent of students was the Terra Nova, taken by 3.26 percent of participants.



Schools have flexibility as to when they administer their exams, and 15 percent of participating students took their exam in the fall months. These scores are less likely to be directly comparable to public school students' tests than are those taken during the time immediately surrounding the public schools' test administration. Table II.2 presents the percentage of students taking the test in the spring, broken down by test taken. It is apparent that the two tests that are least likely to coincide with the public school testing

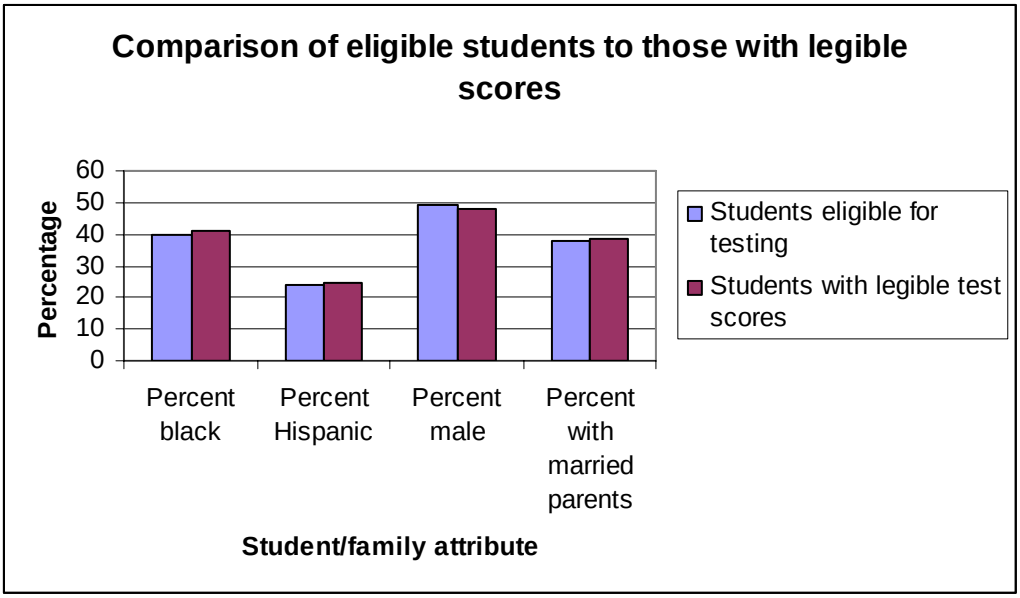
regime are the PSAT/NMSQT and the Iowa Test of Basic Skills. The latter case is driven strongly by Florida Catholic schools' uniform assessment of students in October using the Iowa Test of Basic Skills. It is inappropriate to directly compare status scores of tests administered in March to tests administered in October, as they likely to have very different purposes. This speaks to the importance of measuring student learning gains rather than levels comparisons, and also indicates that it would be useful to conduct a fall-spring concordance study if at all possible.

Representativeness of the sample of students with legible test scores

As mentioned above, there exist 7,764 potential test score records among CTC Scholarship participants in 2006-07, but for various reasons one only observes test scores for 7,067 of these students. The logical next step is to determine the degree to which the students for whom legible test score records are observed are representative of the total set of students who were enrolled in the program at the time of test administration.

Table II.3 compares data collected from CTC Scholarship applications for the 7,067 students with legible test scores as compared with the full set of 7,764 students with potential test score records. These two groups of students are compared along a variety of lines reported on the scholarship applications -- student race/ethnicity, sex, family income, household size and parental marital status. It is clear that the set of students for with legible test score records closely resembles the full set of potential test score records. In no case is the set of students with test scores statistically distinct from the overall characteristics of the potential sample. One can therefore conclude that the

sample of students with legible test score records is a sufficiently representative sample of population of students for whom one could expect test scores.



Because all public school students take the Stanford Achievement Test – also known as the FCAT Norm-Referenced Test -- in March, the most natural direct comparison with public school students would be to study the set of CTC Scholarship participants who attend schools that administer the Stanford Achievement Test in the spring. The final column of Table II.3 demonstrates that the basic attributes of this subset of the overall CTC Scholarship recipient population is observationally equivalent to the overall set of recipients. Therefore, while this group should not form the basis for all comparisons (after all, Catholic school students in Florida are excluded from this group) it is very helpful to observe that this subset of the population is so similar to the overall population of students.

Table II.1: Test administration and score reporting, 2006-07 baseline

Test score reporting status	Number of students	Percent of total
TOTAL NUMBER OF PARTICIPATING STUDENTS	9721	100%
NOT ENROLLED AT TIME OF TESTING	1892	19.46%
Student left program prior to school test administration	1491	15.34%
Student arrived in school after test administration	310	3.19%
Student changed schools midyear between test windows	91	0.94%
INELIGIBLE FOR TESTING	65	0.67%
Student not in grades 3-10	26	0.26%
Student certified to be disabled	39	0.40%
TOTAL NUMBER OF ENROLLED STUDENTS ELIGIBLE FOR TESTING	7764	79.87%
SCHOOL CLOSED / STUDENT CHANGED SCHOOL	124	1.28%
School closed or left program, no tests received	86	0.88%
Student left program, but timing of test unknown	8	0.08%
Student changed schools, no tests received	30	0.32%
USABLE TEST SCORE NOT RECEIVED	573	5.89%
Student certified to be sick/absent during testing period	329	3.38%
Incomplete test reporting	53	0.55%
Test scores reported but school copy is illegible	191	1.96%
TOTAL NUMBER OF LEGIBLE SCORES RECEIVED	7067	72.70%

Table II.2: Distribution of tests administered to Corporate Tax Credit scholarship students

Test	Percentage of total tests	Percent of this test administered in spring months
Stanford Achievement Test	65.78%	98.87%
Iowa Test of Basic Skills	22.70%	45.98%
Florida Comprehensive Assessment Test	4.71%	100%
Terra Nova	3.25%	99.07%
PSAT/NMSQT	1.08%	36.49%
ACT/PLAN	0.74%	86.54%
Metropolitan Achievement Test	0.65%	100%
Basic Achievement Skills Inventory	0.62%	100%
Other tests	0.46%	75%

Table II.3: Comparing students who took tests with legible scores to the full potential population

Student attribute	Students enrolled in program at time of testing (7,764 students)	Students with legible tests in 2006-07 baseline (7,067 students)	Students with legible tests taking school-administered Stanford in spring (4,649 students)
Student is black	39.6%	41.1%	41.8%
Student is Hispanic	23.8%	24.7%	24.3%
Student is white	23.1%	22.7%	22.9%
Student is male	49.1%	48.3%	47.9%
Family income as percent of poverty	124	124	124
Average household size	4.18	4.28	4.28
Parents married	38.2%	38.5%	38.5%

III. Statistics of the 2006-07 baseline

The most natural way to compare students across tests and grades is to measure all students in terms of their national percentile ranking. Table III.1 presents the basic distribution of national percentile rankings among CTC Scholarship students participating in the program in 2006-07, as well as those for only those students attending schools that administered the Stanford Achievement Test in spring 2007. It is apparent that reading and mathematics test scores are normally distributed in this population.

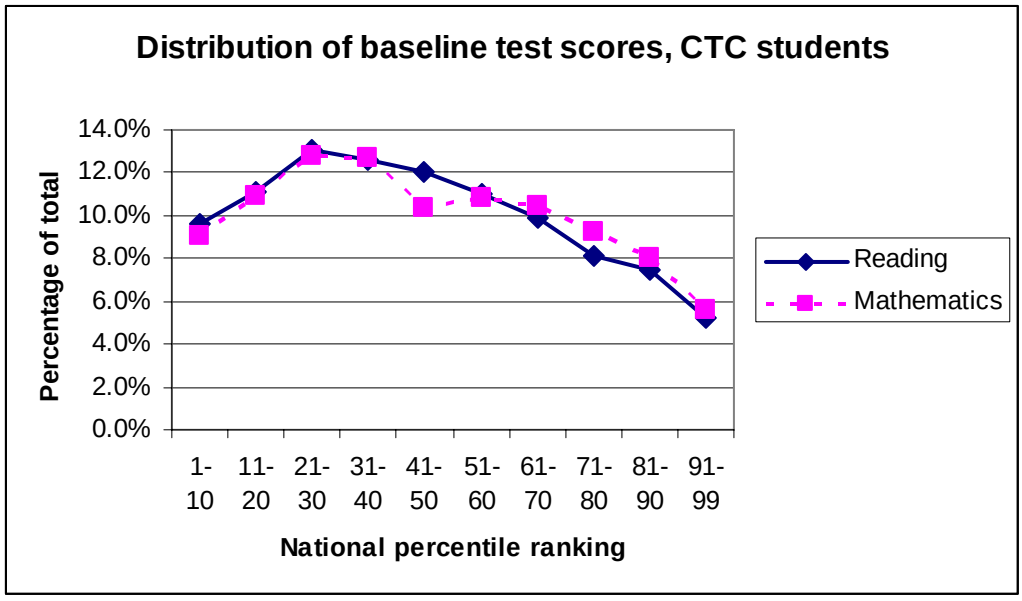
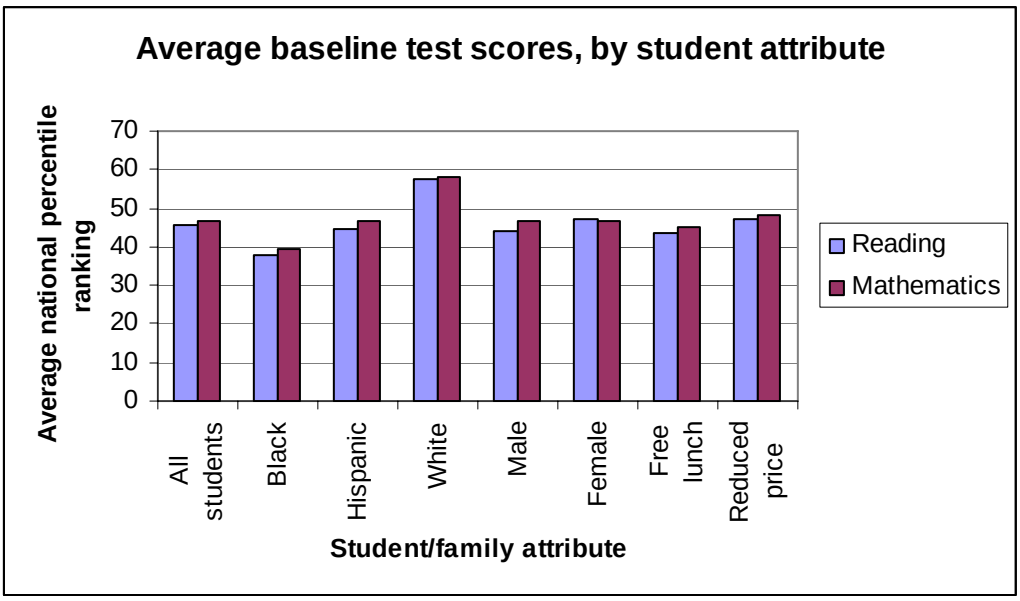


Table III.2 presents average norm referenced test scores for various subsets of the CTC Scholarship recipient population, stratified by race, sex, income, parental marital status and household size. One observes, for instance, that the typical student in the program scored at the 45.48th percentile in reading and at the 46.72th percentile in mathematics. White participants tend to score better than do minority participants, girls tend to perform better than do boys, children with married parents tend to score better than do children with unmarried parents, children from larger families tend to score better than do children from smaller families, and relatively high-income families tend to score

better than do relatively low-income families.[2] Students in schools that administer the Stanford test in the spring months mirror those in the CTC Scholarship population in general, although they tend to perform slightly worse (around one national percentile) in reading, though not in mathematics.



Because the evaluation is designed to consider 2006-07 as the baseline test year, it is important to determine the degree to which the set of students who took 2006-07 tests remaining in the program in the 2007-08 school year resemble the full set of students who took 2006-07 tests in the private sector. As can be seen in Table III.3, the distribution of student test scores for 2006-07 CTC Scholarship students present in private schools in the program in 2007-08 is highly comparable to the distribution of student test scores for 2006-07 CTC Scholarship students in general. This finding reduces the concern that non-random program attrition might interfere with program evaluation.

[2] The income threshold for free lunch participation is 130 percent of the federally-determined poverty level of income for a particular household size. Students in households with income less than 185 percent of the poverty level qualify for reduced-price lunches.

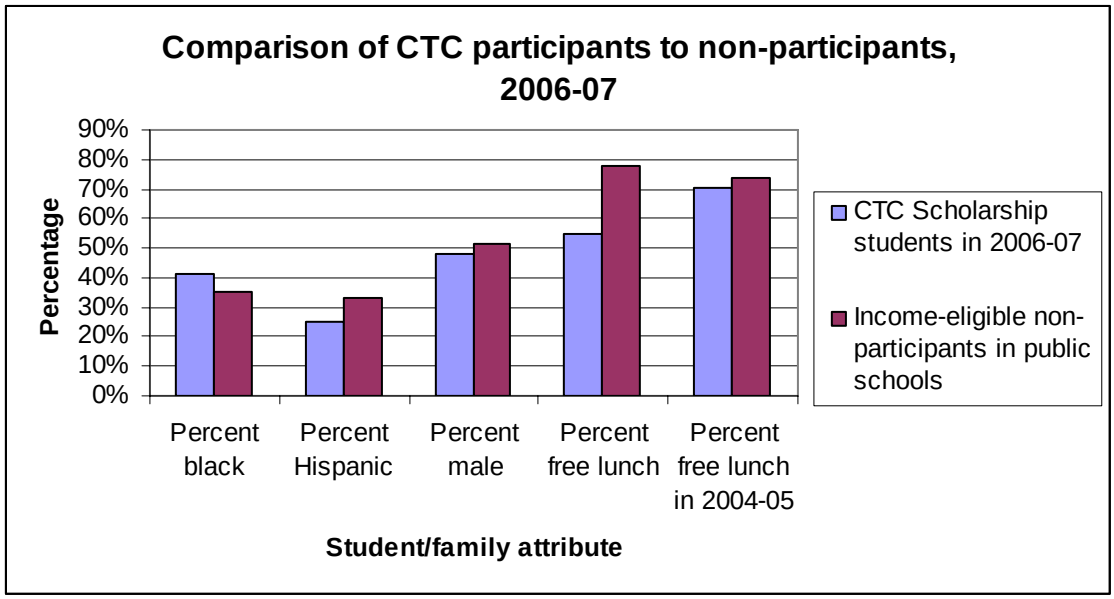
Comparison of public school and CTC Scholarship students' baseline scores

One ultimate purpose of this evaluation is to compare the relative year-to-year gains in the test score of CTC scholarship students to those of comparable public school students. The natural starting-off point, therefore, is to directly compare the performance of all CTC scholarship students to all income-eligible public school students in the baseline year. Such a comparison does not measure the differences in performance of the two sets of schools; doing so would require the comparison of test score gains from 2006-07 to 2007-08. Rather, differences in test scores in the baseline year could reflect either differences in school performance or differences in student and family attributes, or both. One major difference between the sets of students is that CTC scholarship students opted to leave the public schools, generally with at least moderate levels of personal expense; to the extent to which this might have been due to student underperformance, underperforming students may be overrepresented in the CTC Scholarship program.

Another major difference is that while all CTC Scholarship students are certified to be low-income, only three percent of free- or reduced-price lunch students' family incomes are audited, so some fraction of the public school comparison population may actually be of higher income than the program allows. It is difficult to gauge the degree to which either of these differences are relevant, hence the need to compare gain scores rather than status scores.

In addition to difficult-to-measure ways in which CTC scholarship students and measured income-eligible public school students differ, the two groups of students differ

in measurable ways as well. For instance, Table III.4 presents the grade distribution of income-eligible test-takers in the public schools as compared with the grade distribution of CTC Scholarship test-takers. It is apparent that the CTC Scholarship students are much more likely to be in the elementary grades and less likely to be in the secondary grades, and especially in high school, than are public school students. This partially reflects the fact that there are more private schools for elementary and middle school students, and partially reflects the relative youth of the CTC Scholarship Program.



The two groups of students differ in other ways, as well. Table III.5 presents some of the mean attributes of the two groups. One observes that CTC Scholarship test-takers are more likely to be black, less likely to be Hispanic and less likely to be male than are income-eligible test-takers in the public schools. However, CTC Scholarship test-takers are less likely to be reported to be free-lunch eligible than are income-eligible public school students. This last comparison could reflect true differences in income distribution or it could reflect misrepresentation of income status on school meal

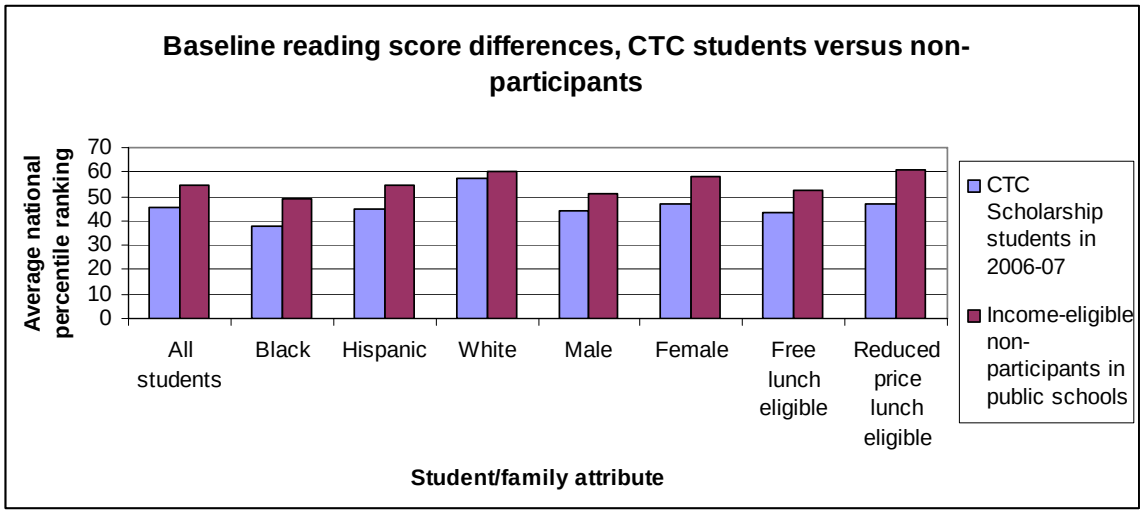
qualification applications, or a combination of both.[3] As a way of gauging the degree to which this might be the case, one can compare the public school records of the two groups of students from a prior year; in that year, both sets of students' incomes were subject to the same reporting regime. It turns out that the largest number of public school student records were found two years prior, in 2004-05. This may be due to the fact that although students have some probability of leaving the state over time, they are relatively likely to remain in the same schooling situation from one year to the next. When this comparison is made, it is apparent that, historically speaking, the two groups of students are more similar on the grounds of reported income: Among CTC Scholarship students found in the public schools in 2004-05, 70.3 percent reported being free-lunch eligible, as compared with 74.1 percent of non-participating eligible students in 2006-07.[4] Such a comparison suggests that either families of participating students experienced a sudden relative uptick in income in the year in which they applied for the program or that families may tend to under-report their incomes on free meal qualification applications. While it is impossible to determine with certainty which of these explanations is the most likely, this discussion makes clear how important it is to identify appropriate comparison groups.

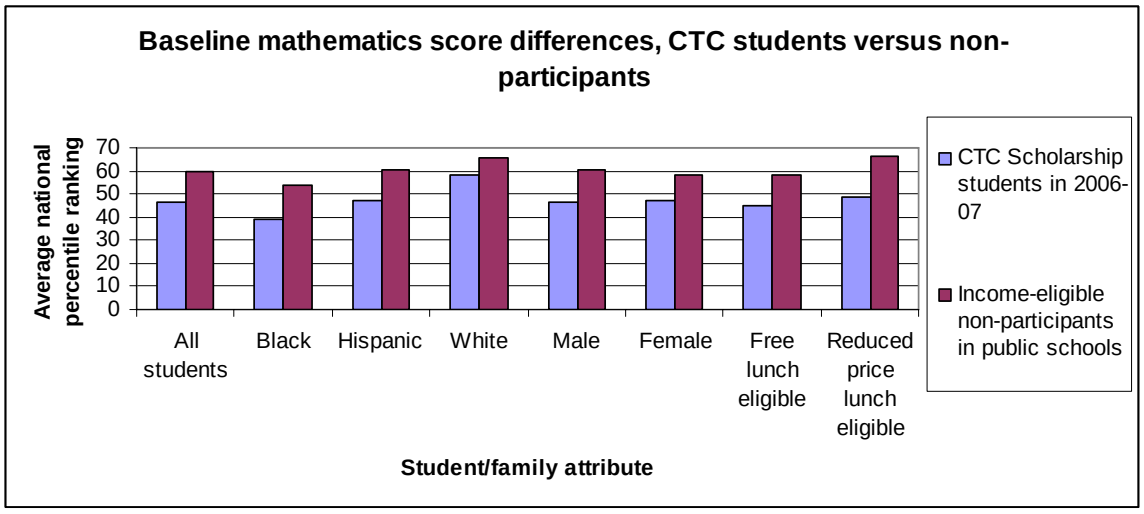
While the previous provisos make it clear that it is inappropriate to draw inference from a comparison of the two groups of students in the baseline year of 2006-07, it is still

[3] The income measured in the CTC Scholarship participant population comes from federal tax returns rather than income reported on school forms and audited at a three percent rate, and is therefore less likely to be misrepresented.

[4] Incidentally, the rate at which test-taking CTC Scholarship students are present in 2004-05 in the Education Data Warehouse is almost as high (86 percent) as the rate at which non-participants are present in 2004-05 (91 percent.)

a useful starting-off point to make such a comparison. Table III.6 presents average 2006-07 norm-referenced test scores for all non-participating free- or reduced-price lunch eligible students in Florida public schools, broken down by subgroup, compared with the same subgroups of CTC Scholarship recipients.[5]





Because the average baseline test scores are sufficiently similar between all CTC Scholarship recipients and those who took a school-administered Stanford Achievement

[5] Several of the student and family attributes compared above, such as parental marital status, household size, and detailed family income (rather than simply free- or reduced-price lunch eligibility) are only observed on scholarship applications and not for the population of income-eligible students in general, so they cannot be used to create subgroups for this comparison.

Test in the spring, for the remainder of this report only the baseline results for all CTC Scholarship students are reported. The comparison test scores for non-participating public school students is the Stanford Achievement Test (the FCAT Norm-Referenced Test) administered in March 2007.

It is obvious from a perusal of Table III.6 that students participating in the CTC Scholarship program average considerably lower on nationally norm-referenced tests than do non-participating public school students who are income-eligible. On average, CTC Scholarship students score nine national percentiles lower on their reading examinations and 13 national percentiles lower on their mathematics exams than do income-eligible non-participants. The differences are particularly pronounced for black students in reading and both black and Hispanic students in mathematics. The differences are larger for females in reading and for males in mathematics, and the differences are larger for students in families with incomes between 130 and 185 percent of the poverty line than for students in families with incomes below 130 percent of the poverty line, the free- and reduced-price lunch thresholds. Again, however, the reader should be cautioned to not draw inference from these comparisons.

Potential treatment and comparison groups

When CTC participants' and non-participants' test score gains from 2006-07 to 2007-08 are to be compared, it is important that the groups being compared are as similar as possible on observed attributes. One possible approach to doing so employs a so-called propensity score matching estimator, in which each CTC Scholarship participant is

paired with a randomly-selected non-participating income-eligible student from the public sector who is identical along a number of observed dimensions (race/ethnicity, baseline test score, sex, and geographical location) and the two sets of students are followed from 2006-07 to 2007-08. While the follow-up report will perform this type of analysis, it is still important to note that such a comparison does not take into account the fact that people are not randomly assigned to participate in the program. Therefore, while a propensity score matching estimator would generate much more balanced comparisons between participants and non-participants than would raw comparisons such as those presented in Table III.6, doing so would by no means guarantee an accurate assessment of the relative effect of CTC Scholarship receipt for participating students.

One potential solution to the problem of non-random selection into the CTC Scholarship program is to compare participating applicants to non-participating applicants. This type of comparison has the distinct advantage of identifying the effects of CTC Scholarship participation based on the performance of individuals who, even if they did not use the scholarship, were at least sufficiently likely to select into the program that they completed an application and paid an application fee. There are numerous ways in which applicant-to-applicant comparisons can be made, and will be made in the second year report, and fully exploring the potential comparisons requires more time than has been available for this baseline report. That said, even in this baseline report a few applicant-to-applicant comparisons can still be made.

The ideal design for a study of the effects of CTC Scholarship Program participation would involve an experiment, in which some students received the scholarship (and were therefore considered to be part of a "treatment group") while others were similar along a number of dimensions but did not receive the scholarship (and were therefore considered to be part of an experimental "control group.") While no experiment was conducted, it is still possible to use the experimental design nomenclature to compare scholarship participants (the treatment group) to similar non-participants (the control – or comparison -- group). The last task of this baseline report is to assess the degree to which treatment and control groups can be constructed from the available data. The results of this exercise are presented in Table III.7.

The top panel of Table III.7 presents a series of potential treatment categories – students who are active in the scholarship program in both the baseline year and the current year; students who are active in the program in the current year but not the baseline year; and all students currently active in the program.[6] Each of these three specifications of the treatment group has advantages and disadvantages. In many regards, the second of these candidate treatment categories – those who are new to the program in 2007-08 – has appeal, because one can interpret any comparisons as a first-year effect of participation. However, it has its drawbacks as well – it is important to understand the effects of program participation past the first year (especially given that first year results can confound transition effects with program effects) and newcomers to the program comprise only about one-third of the program participants. Therefore, the follow-up

[6] Of the 7,067 students with reported legible test scores in the 2006-07 round of data collection, 5,223, or 74 percent, were still participating in the CTC Scholarship Program at the time of writing.

report will investigate student learning gains for each of these groups of program participants.

An issue in finding appropriate comparison groups is that the baseline test scores of the potential characterizations of the treatment group are quite different. Newcomers to the program average around ten national percentiles higher reading and mathematics scores than do returning students in the program. Along other dimensions, however, the two groups of 2007-08 program participants appear to be similar. The fact that these two groups differ substantially in baseline test scores, however, indicates that much care will be needed in identifying separate comparison groups for these two different treatment groups.

There are several ways in which application data may be used to construct potential comparison groups.[7] One possible approach involves the fact that scholarship applicants are frequently ruled to be ineligible typically either due to parental income being too high, a student not having attended a Florida public school for the full year prior to enrollment in the program, or (in rare instances) fraudulent applications. In 442 cases, students ruled ineligible in 2007-08 have observed 2006-07 test scores. One observes that the 2006-07 test scores for ineligible students who were not in the program in 2006-07 (the fifth row of the table) are similar to those students who were new to the program in 2007-08 (the second row of the table). However, some of the demographic attributes differ between the two groups, most notably income (the potential treatment

[7] Because of time limitations, my understanding of the specifics of the application data is imperfect. Therefore, the work constructing potential comparison groups is extremely preliminary and subject to change before the final report.

group has income averaging just 119 percent of poverty, while the potential comparison group has income averaging 163 percent of poverty.) The percentage Hispanic in the potential comparison group is also five percentage points higher than in the potential treatment group.

It is, however, possible to limit the income range of the potential treatment group to more closely adhere to the potential comparison group. In results not shown herein, if one limits the set of newcomers to the program to be those with incomes above 130 percent of the poverty line, the typical income in that restricted sample is 158 percent of poverty (compared with 163 percent in the comparison group), and test scores average 55.22 points for reading (compared with 54.97) and 58.22 points for mathematics (compared with 58.96). Hence, further restricting the treatment group for the purposes of analysis has the benefit of increasing believability in the outcomes. On the other hand, it leads to a further restricted sample. To deal with these concerns, the follow-up study will estimate a regression-discontinuity model that would make these comparisons more concretely; doing so would require the researcher to have a better understanding of the reasons why families apply for scholarships for which they are unqualified due to income.[8] A regression-discontinuity model has particular appeal because it allows the researcher to take into account the fact that test scores tend to correlate positively with income, making higher-income students a less desirable comparison group for lower-income students.

[8] This spring the Independent Research Organization will be conducting focus groups of applicants and potential applicants. The Project Director expects that these focus groups will shed light on this and other important questions that will inform the second year of this study.

Similar comparisons can be made for students who were in the program in 2006-07 and remain in the program vis-à-vis those in the program in 2006-07 but ruled ineligible for continuation in 2007-08, and the baseline test scores for these two groups are reasonably similar, especially in the case of mathematics. Even more of these students, however, are ineligible due to income, and the average income level of this potential comparison group is 200 percent of the poverty line. That said, these students are probably less likely to be considered ineligible due to confusion about program eligibility, so the students who become marginally ineligible may still make a reasonable comparison group for the students who remain marginally eligible.

A different type of comparison group comes from the set of students who are voluntarily not using the scholarship. The third panel of Table III.7 presents some basic information about these students. The test scores and income levels of these students are reasonably similar to those of the potential treatment groups, although the percentage black in the not-using categories tends to be somewhat higher than in the using categories. The most fundamental problem with using non-users as a comparison for users is that there may be systematic reasons for the lack of use. The follow-up report will present the results of an investigation into whether there are spatial differences (e.g., urban/rural or in different parts of the state) that help to determine the degree of scholarship usage contingent on offer. The degree to which external reasons why some students are more likely to take the scholarship than are others – for instance, differences in access to private school spaces or differences in private school tuition – can be

identified will determine whether non-users might make an acceptable comparison group in the final report.

Table III.1: Distribution of national percentile rankings for participants in the CTC Scholarship program, 2006-07

National percentile	Reading		Mathematics	
	Students with legible tests in 2006-07 baseline	Students with legible tests taking school-administered Stanford in spring	Students with legible tests in 2006-07 baseline	Students with legible tests taking school-administered Stanford in spring
1-10	9.60%	10.88%	9.11%	9.39%
11-20	11.05%	11.31%	10.86%	10.97%
21-30	13.11%	14.11%	12.83%	13.00%
31-40	12.58%	12.50%	12.74%	12.30%
41-50	12.04%	11.78%	10.39%	9.79%
51-60	11.02%	10.54%	10.77%	10.75%
61-70	9.85%	9.61%	10.49%	10.64%
71-80	8.07%	7.47%	9.23%	9.95%
81-90	7.51%	7.20%	7.96%	8.19%
91-99	5.17%	4.60%	5.62%	5.02%

Table III.2: Average national percentile rankings for participants in the CTC Scholarship program, 2006-07, for students of different background characteristics

Characteristic	Reading		Mathematics	
	Students with legible tests in 2006-07 baseline	Students with legible tests taking school-administered Stanford in spring	Students with legible tests in 2006-07 baseline	Students with legible tests taking school-administered Stanford in spring
All students	45.48	43.98	46.72	46.64
Black	37.91	36.62	39.18	39.46
Hispanic	44.59	42.59	46.79	46.23
White	57.58	56.64	58.31	58.32
Male	43.88	42.35	46.59	45.88
Female	46.97	45.48	46.84	47.33
Household size <=4	44.45	42.68	45.29	45.22
Household size >=5	47.13	45.99	49.07	48.90
Free lunch eligible	43.73	42.15	45.19	44.87
Reduced price lunch eligible	47.15	45.95	48.36	48.60
Parents married	49.90	48.50	51.36	51.41
Parents unmarried	42.77	41.18	43.87	43.68

Table III.3: Average national percentile rankings for participants in the CTC Scholarship program, 2006-07, for students of different background characteristics: Students continuing in the program in 2007-08

Characteristic	Reading		Mathematics	
	Students with legible tests in 2006-07 baseline	Students with legible tests taking school-administered Stanford in spring	Students with legible tests in 2006-07 baseline	Students with legible tests taking school-administered Stanford in spring
All students	45.32	43.85	46.84	46.60
Black	37.41	36.62	39.05	39.66
Hispanic	43.96	42.16	46.06	45.61
White	58.55	57.19	59.36	58.64
Male	43.89	42.42	46.66	46.05
Female	46.64	45.16	47.01	47.11
Household size <=4	44.20	42.69	45.36	45.36
Household size >=5	47.13	45.72	49.22	48.62
Free lunch eligible	43.45	41.40	45.14	44.04
Reduced price lunch eligible	47.12	45.92	48.55	48.76
Parents married	49.68	48.17	51.16	51.04
Parents unmarried	42.48	41.09	44.03	43.77

Table III.4: Grade distribution in test-taking grades of CTC Scholarship test-takers and income-eligible public school test-takers

Grade	CTC Scholarship students in 2006-07	Income-eligible non-participants in public schools
3	17.30%	15.10%
4	14.97%	14.09%
5	20.53%	13.45%
6	14.08%	13.61%
7	10.26%	11.90%
8	10.40%	12.60%
9	7.23%	10.67%
10	5.22%	8.49%

Table III.5: Differences in other observed attributes of CTC participants and non-participants

Attribute	CTC Scholarship students in 2006-07	Income-eligible non-participants in public schools
Percent black	41.07%	34.86%
Percent Hispanic	24.73%	32.97%
Percent male	48.26%	51.42%
Percent free lunch	54.84%	77.70%
Percent free lunch in 2004-05 (if observed in public school)	70.30%	74.09%

Table III.6: Average national percentile rankings for participants in the CTC Scholarship program, 2006-07, for students of different background characteristics, as compared to the population of non-participating income-eligible students in the public schools

Characteristic	Reading		Mathematics	
	CTC Scholarship students in 2006-07	Income-eligible non-participants in public schools	CTC Scholarship students in 2006-07	Income-eligible non-participants in public schools
All students	45.48	54.54	46.72	59.77
Black	37.91	49.17	39.18	53.49
Hispanic	44.59	54.41	46.79	60.60
White	57.58	60.05	58.31	65.24
Male	43.88	50.78	46.59	60.55
Female	46.97	58.43	46.84	57.95
Free lunch eligible	43.73	52.68	45.19	57.95
Reduced price lunch eligible	47.15	60.99	48.36	66.12

Table III.7: Comparison of 2006-07 baseline test scores and student attributes for different groups of students in 2007-08

Group of students	Average reading percentile	Average math percentile	Percent black	Percent Hispanic	Income as pct. of poverty	Average family size	Percent parents married
POTENTIAL TREATMENT CATEGORIES							
In program in 06-07 and 07-08	45.34	46.86	39.42	24.30	122	4.22	40.35
In program in 07-08, not in 06-07	54.80	57.10	38.00	24.02	119	4.10	37.48
All students in program in 07-08	48.37	50.13	38.82	24.19	122	4.17	39.15
POTENTIAL COMPARISON CATEGORIES							
In program in 06-07, ineligible in 07-08	48.68	46.99	43.18	26.52	200	4.22	44.53
Not in program in 06-07, ineligible in 07-08	54.97	58.96	39.04	29.21	163	4.18	35.67
Ineligible in 07-08	53.24	55.66	40.17	28.48	173	4.19	38.13
In program in 06-07, not using in 07-08	45.44	46.53	44.29	21.22	122	4.41	36.29
Not in program in 06-07, not using in 07-08	51.45	56.84	44.67	25.16	121	4.00	26.42
Scholarship awarded, not using in 07-08	47.76	50.50	44.43	22.77	121	4.25	32.48

IV. Conclusion

The principal purposes of this report are twofold. The first purpose is to determine the degree to which private schools complied with the testing requirements in the 2006-07 academic year, the first year in which participating schools were required to test CTC Scholarship students, and to gauge whether 2006-07 could make an appropriate baseline for analysis. The evidence suggests that participating private schools universally complied with the program. Of the schools that continue to participate in the program, there were only reporting errors from six schools, and most of the reporting problems stemmed from an apparent misunderstanding of the reporting requirements of the program. The largest issue with the baseline test collection is not one of compliance with the testing requirement per se, but rather the number of illegible copies of score sheets – just under two percent of the students in the program. That said, the evidence suggests that the illegible scores are representative of the overall population of participating students, indicating that the 2006-07 school year makes for a solid baseline year for studying the effects of the Corporate Tax Credit Scholarship program.

The second primary purpose of this report is to provide baseline statistics on average test performance, overall and by subgroup, among CTC Scholarship participants, and to describe the performance differences and characteristics across various potential treatment and control groups. Participant test scores in the CTC Scholarship program run the gamut from very high performance to very low performance, and average lower test scores than in the public sector. However, this test score gap cannot be interpreted as a difference in performance between public and CTC Scholarship schools, as the two sets

of schools attract different groups of students. The follow-up report that incorporates student learning gains between 2006-07 and 2007-08 will provide very important evidence concerning the potential success of the CTC Scholarship program in improving the performance of Florida's low-income population.