Richardson *MathForward* Project
Second Year Final Report: Math TAKS Results

Walter Stroup, Ed.D., Vinh Pham
and Celeste Alexander Ph.D
The University of Texas at Austin

Introduction

The Richardson Independent School District in Richardson Texas has implemented a novel program to improve mathematical skills for some 7th, 8th, and 9th graders called *MathForward*. With the help of new technology and innovative assessments students are able to communicate their mathematical thinking and then receive immediate feedback regarding their mathematical knowledge.

Earlier pilot results indicate several components of the intervention are crucial to the success of the intervention. The key components of *MathForward* include: extended learning time, use of technology to motivate and enhance learning opportunities, provision of common, aligned assessments, increased teacher content knowledge, and development of high expectations for all students.

The teachers involved in the *MathForward* program received specialized professional development training. Students participate in a 100-minute mathematics class that focuses on enhancing mathematical understanding through the use of TI Navigator™ system and advanced graphing calculators, in-classroom networks and daily problem solving. The students also participated in daily lessons where they must communicate solutions, apply content, and connect mathematical models to abstract concepts. The technology allows teachers to monitor, on a screen, each student's progress as concepts are taught and problems are worked. This allows for immediate feedback and opportunities for intervention.

This report is focused on both the analyses from 2005-06 and the continuation into this year (2006-07) for the intervention study. To ensure a clear and concise understanding, a consistent vocabulary bridging the two years was needed. The following definitions are used for the remainder of this report:

Type 1 – Students who failed the previous year's TAKS
Type 2 – Students who passed the previous year's TAKS
Study – Students who are in the study intervention (2005-06 only students that failed the previous TAKS, *2006-07 inclusive* – combined students, some who passed and some who failed the previous TAKS)
Comparison – Students who are at the comparison campus*
Control 1 – Non-study students who are taking other math classes with study teacher
Control 2 – Students who are taking classes with non-study teachers at study schools
Control 3 – Students who are taking classes at non-study schools

*Forest Meadows was the comparison campus and was chosen for similar demographics to the study school. This was for year one only.

Methodology

Student Demographic Information

Data provided by the district includes indicators for student ethnicity and whether student is classified as economically disadvantaged this is referred to as "SES" in this report. There were no other indicators such as classification as Limited English Proficient (LEP) or participation in Gifted and Talented classes. Students included in the analyses were required to have both a 2006 and 2007 math TAKS score (so change could be assessed). This means that highly mobile students tend to be excluded from the analysis. Students were 7^{th}, 8^{th}, and 9^{th} graders in the 2006-07 school year in regular math classes. The variable "SES" is the indicator for whether a student is participating in the federal Free and Reduced Priced Lunch (FRPL) program.

Regression Discontinuity

Before continuing, a description of regression discontinuity and its usefulness in *need-based* programs is necessary. Regression discontinuity studies rely on the hypothesis that observations will have a different pattern at a pre-defined point on a continuum (Cook and Shadish, 1994, Shadish, Cook, Campbell, 2001). Or in other words, in the absence of the treatment program, the pre-post relationship of the groups would be equivalent (Trochim, 2006). The regression discontinuity is basically a pretest-posttest program-comparison group design. This type of design is appropriate when educators want to target a program to students who need intervention the most (Trochim, 2006). Assignment of participants to a particular treatment or programs is based on a cutoff point. Because we know that the treatment and control group means differ since the group assignment is based on a pretest score, we can estimate the treatment effect by the size of the projected discontinuity (jump or change) at the cutoff. For the regression discontinuity design, if there are possible interactions or nonlinearities, it is very important to include these in the model. It is better to over-fit a model because over-fitting still yields unbiased coefficients, although it decreases power. Over specification assures that all necessary terms have been included even at the expense of unnecessary ones.

For the 2005-06 *MathForward* intervention, a "pull out" approach was implemented for the 2006 academic year. Students below the passing score for the Texas Assessment in Knowledge and Skills for mathematics in 2005 were assigned to separate classes and subject to a complex, but well-defined, set of interventions that constitute the the *MathForward* program. When scores are plotted using Normal Curve Equivalent (NCE), significant positive discontinuities were observed in comparing the 2005 TAKS results with the 2006 TAKS mathematics scores. These results can be seen to provide strong evidence for the causal validity of this treatment. Results from ordinary least squares analyses of the data set also showed significance as well as effects of comparable size to those resulting from the RDD analyses. OLS analyses also provide additional complementary information about the relative significance of factors like ethnicity and/or socio-economic status.

Based on the judged success of the first year's intervention, the scope of the implementation was broadened in the 2006-07 school year.

Inclusive 2006-2007 Regression Discontinuity Design

As part of the scaling-up of the *MathForward* intervention, students below the cut score were included in the same classes as students above the cut-off score. Methodologically this moves the use of regression discontinuity techniques beyond the "pull out" interventions with which it is typically associated. In medical contexts, interventions are administered to a treatment group that is, by definition, independent of the population above the cut-off score (the non-treatment group). Also, grade 9 students were added in the intervention for this second year. When transferred to educational contexts, the methodological artifact of an experimental treatment taking place in a self-contained intervention (e.g., a certain drug or therapy being administered in physical isolation to individuals in a treatment group) is often maintained by placing students below a cut-score in a self-contained, or "pull out", educational program. This was the model used for the first year of the *MathForward* study (2005-06) and the results of using RDD in this standard way are reported below.

Mathematically, however, the only significant requirement beyond the use of a well-defined cut score for deciding who is in the treatment is that the student population of the treatment group is independent, or orthogonal, to the population of students above the cut score. For year two of the study, being in classes with students above the passing score is considered part of the overall treatment for the students below the passing score. This "inclusive" model for the year-two (2006-07) classroom implementation of the intervention is matched in this report to the use of an "inclusive" regression discontinuity design. This inclusive RDD does maintain the required mathematical independence (orthogonality) of the comparison group.

Normal Curve Equivalent

All scores were derived by converting each student's TAKS scale score into a NCE (Normal Curve Equivalent) score using that year's TAKS score frequency distribution as reported by the Texas Education Agency (2007) for the student's grade level. This approach allowed for arithmetic manipulation of the scores in a way that is consistent and more transparent. The intervals between the scores are now equivalent in a way that would not be the case using the scale scores. To ensure standardization of the scores relative to the cutoff score for each year and grade, the NCE scores were further transformed by subtracting the cutoff NCE score for the grade level from the students' NCE scores. All transformed NCE scores are labeled as TNCE. A TNCE score of 0 would correspond to the cut score for passing the TAKS. Another refinement is that the analyses largely omitted, except where noted, data from students who were in pre-AP classes. These students represented a distinct and skewed subset of the population not comparable to the sample of the population for which the intervention was implemented.

These changes account for the relatively minor differences in the reported results from last year's analysis.

Results

Descriptive Statistics

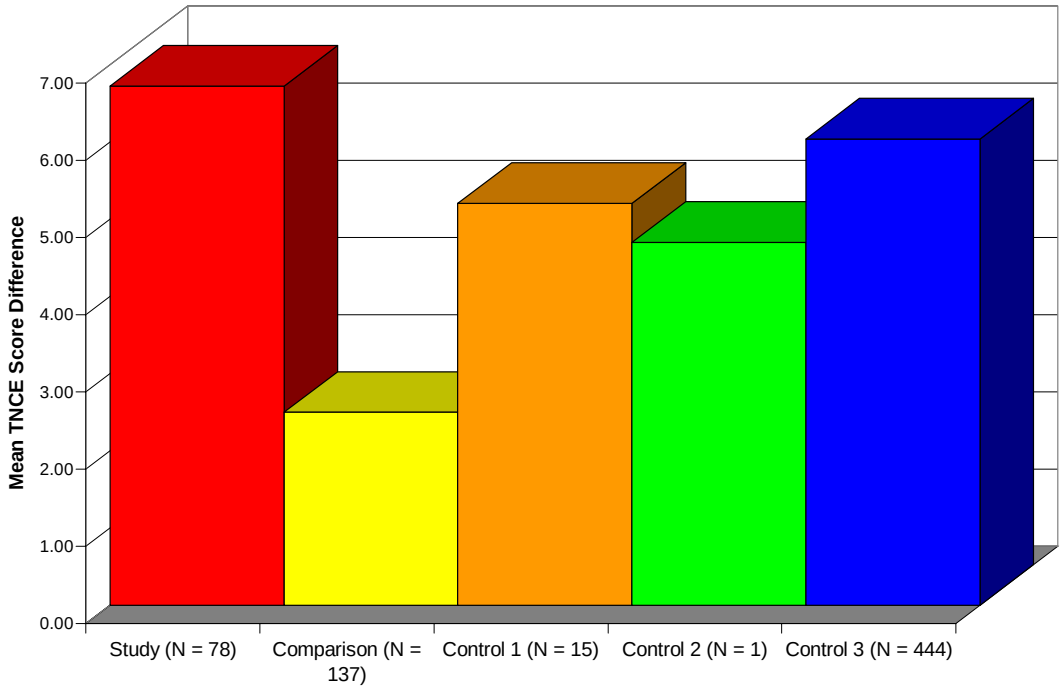
Table 1 shows the descriptive statistics for 2006's data based on the convention set forth above.

Table 1. 2006 Statistic Descriptives

		Minority	Low SES	Passed 2006	2005 TAKS TNCE	2006 TAKS TNCE	Change in TNCE
Type 1	Study (N = 78)	95%	65%	33%	-13.04	-6.32	6.72
	Comparison (N = 137)	96%	81%	19%	-13.97	-11.47	2.50
	Control 1 (N = 15)	87%	67%	13%	-17.94	-12.73	5.21
	Control 2 (N = 1)	100%	100%	0%	-13.30	-8.60	4.70
	Control 3 (N = 444)	73%	63%	37%	-11.26	-5.22	6.04
Type 2	Study (N = 0)	N/A	N/A	N/A	N/A	N/A	N/A
	Comparison (N = 137)	85%	69%	72%	14.97	8.53	-6.44
	Control 1 (N = 105)	72%	47%	77%	14.91	9.35	-5.56
	Control 2 (N = 1)	100%	100%	0%	6.90	-10.70	-17.60
	Control 3 (N = 1309)	52%	45%	88%	19.78	15.52	-4.26

.

Chart 1. Mean TNCE change for different Type 1 students between 2005 and 2006



For 2006, a Control 2 group was added for the sake of completeness and year-to-year

consistency. However with an N of only 1, this value cannot be considered generalized. Students in the study last year showed 6.72 mean TNCE point increase. This is the largest increase in all the Type 1 students as seen in Chart 1. When interpreted on a 100 point scale, this increase is substantial. However, all of the other groups also showed positive gain, though not as much. The Control 3 improvement almost matches that of the Study group. Further analysis will make the contrasts more apparent and serve to underscore the limitations of relying on purely descriptive analyses.

The results for 2007 are largely similar but with some notable variations as shown in Table 2. Data points identified as outliers by SPSS were removed from further analysis. Outliers were defined as any data point that had a standardized residual greater than 3 standard deviations from the predicted score using a linear regression of the data. It can be observed from the results that there was an even larger mean TNCE increase with Type 1 Study students (~8.5 points).

Table 2. 2007 Statistic Descriptives

		Minority	Low SES	Passed 2006	2006 TAKS TNCE	2007 TAKS TNCE	Change in TNCE
Type 1	Study (N = 236)	88%	67%	48%	-11.53	-3.04	8.49
	Control 1 (N = 32)	88%	72%	31%	-15.58	-10.41	5.17
	Control 2 (N = 188)	84%	66%	30%	-14.91	-7.71	7.20
	Control 3 (N = 257)	75%	62%	34%	-10.79	-7.87	2.92
Type 2	Study (N = 134)	75%	63%	86%	6.96	12.36	5.40
	Control 1 (N = 38)	87%	45%	84%	15.33	11.31	-4.02
	Control 2 (N = 584)	54%	42%	91%	18.3	16.02	-2.28
	Control 3 (N = 794)	49%	40%	85%	16.81	13.06	-3.75

As shown in Chart 2, all Type 1 students again showed positive growth with the result for Control 2 students now rivaling that of the Study students. Why Control 2 had similar growth as Study students warrants further investigating. Were there any other intervention going on at the study schools? This was not believed to be a contamination effect since Control 1 students, taught by the study teachers and who would be more likely to exhibit contamination effects than Control 2 students, did not show a similar magnitude of growth. Moreover, a significant element in the intervention was the use of specific network technology (TI-Navigator™) and because this tool was not available to students outside of the study, contamination effects are even less likely to account for the somewhat similar gains by Control 2 students.

Chart 2. Mean TNCE change for different groups of Type 1 students between 2006 and 2007

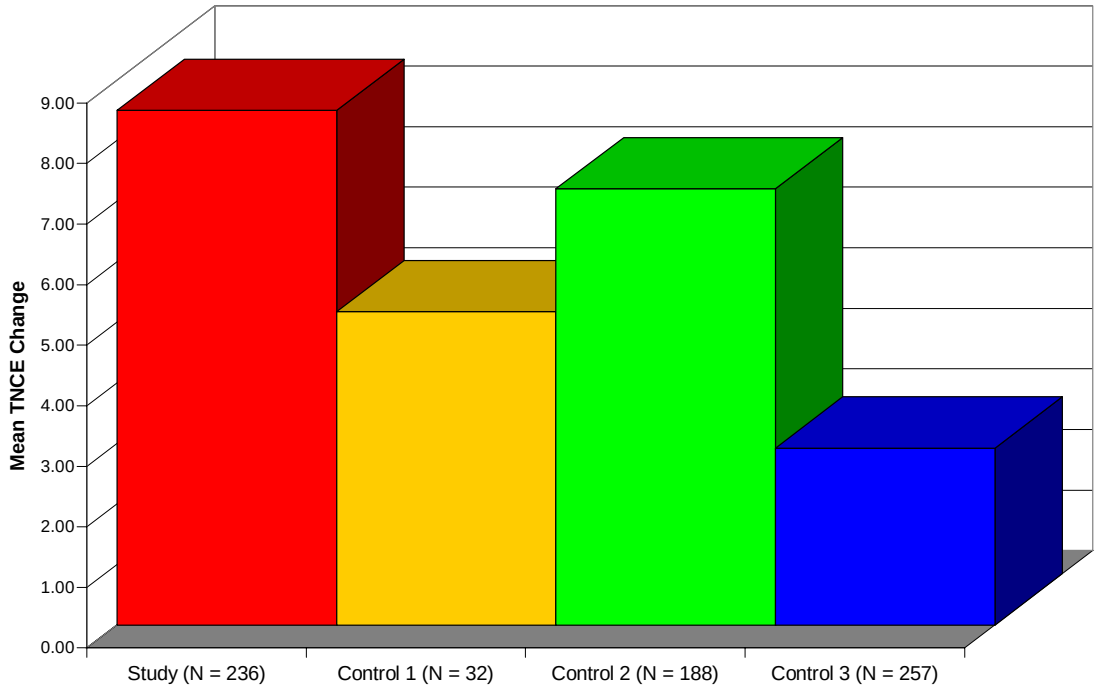
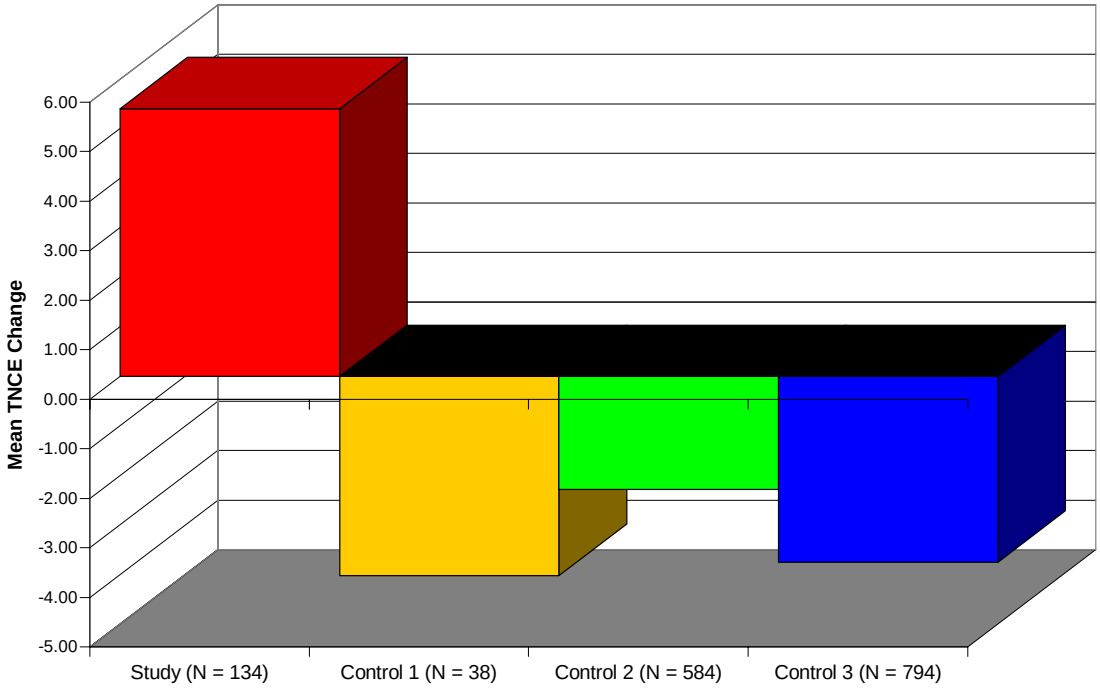


Chart 3. Mean TNCE change for different groups of Type 2 students between 2006 and 2007



To further highlight what may be happening in the classrooms where the Control 2 students made gains, it is revealing to plot the mean TNCE changes for all the Type 2 (passing) students. The *only* Type 2 students to have shown any positive gains were the Study students as depicted in Chart 3. This suggests that while the intervention was

clearly beneficial to *both* Type 1 and Type 2 students, whatever methods might have been employed to improve Type 1 student scores in the Control 2 classrooms these approaches only improved the results on the Type 1 students (and even then, not by as much as the intervention did for Type 1 Study students). Type 2 students were negatively impacted in the Control 2 classes. Even as the intervention was motivated primarily by the desire to improve the results of Type 1 students, it is important nonetheless to also report results for the passing students. In light of the gain for both Type 1 and Type 2 students in the intervention, additional effort might be expended in understanding the intervention efforts for students already achieving at the passing level. The pull-out approach used in 2006 means there could be no data available for this group in that year.

Ordinary Least Squares (OLS) Analyses

The following ordinary least squares regressions use transformed NCE scores. Again, using TNCE allows for comparison of the students across years and grades at the cutoff point. This allows for a more accurate analysis of the effects of the study intervention and its effects on students' passing rates on the TAKS. It is important to note that OLS regression analysis uses the complete data set of both Type 1 and Type 2 students of each group under consideration. This is especially significant for year two results since year one had no Type 2 Study students.

Table 3. OLS Regression for 2006 Study and Comparison students.

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
2005 TNCE Score	0.653	0.037	0.718	17.667	0.000
Study Treatment	4.130	1.558	0.108	2.650	0.008
Minority	-4.130	2.259	-0.070	-1.774	0.077
SES	-0.441	1.427	-0.012	-0.309	0.757

Dependent variable – 2006 TNCE
R^2 = 0.495, Durbin-Watson = 1.373, N = 352

Using an ordinary least squares analysis, the treatment effects were significant when compared to the comparison students as shown in Table 3. Minority and SES were not significant, though minority was trending towards significance. Study students were predicted to have ~ 4 NCE points higher than the comparison students.

Table 4. OLS Regression for 2006 Study and Control 1 students

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
2005 TNCE Score	0.668	0.062	0.763	10.815	0.000
Study Treatment	2.972	2.152	0.095	1.381	0.169
Minority	0.976	2.308	-0.024	0.423	0.673
SES	-0.992	1.720	-0.032	-0.577	0.565

Dependent variable – 2006 TNCE
R^2 = 0.496, Durbin-Watson = 1.342, N = 198

Table 4 shows that the Control 1 and study students were not statistically distinct. This lack of statistical significance could be accounted for by the smaller N value of Type 1 Study and Control 1 students as well as by the absence of Type 2 Study students causing statistical artifacts to appear in the OLS results. Even with a lack of significance, the OLS results do show that study students tended to have almost 3 TNCE points higher than Control 1 students and that study treatment effects can explain almost 10% of the variance seen in the data. A larger N value might well have allowed the study treatment to become significant. This larger study N was achieved in year two and is discussed below. Minority status and SES were not significant. This was likely a result of insufficient variation in the data due to the relative homogeneity of the sample used. For these variables, a small N value could have, once again, made getting to significance more difficult.

Table 5. OLS Regression for 2006 Study and Control 3 students.

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
2005 TNCE Score	0.611	0.016	0.670	38.667	0.000
Study Treatment	0.606	1.443	0.007	0.420	0.674
Minority	-4.912	0.673	-0.139	-7.301	0.000
SES	-0.611	0.649	-0.018	-0.942	0.347

Dependent variable – 2006 TNCE
R^2 = 0.524, Durbin-Watson = 0.680, N = 1831

When compared to the other schools in the district for 2006, the data is inconclusive (Table 5). Study effects were insignificant. This, again, is believed to be due to the small population in the Study (a situation that changes for 2007). Minority status proved to be significant while low SES was not. While these two variables are often co-incident, it is interesting to see them different from each other in these results. Both were found to be negative in their impact on scores.

Table 6. OLS Regression for 2007 Study and Control 1 students.

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
2006 TNCE Score	0.728	0.041	0.650	17.803	0.000
Study Treatment	5.443	1.500	0.132	3.629	0.000
Minority	-4.557	1.567	-0.110	-2.908	0.004
SES	-.790	1.194	-0.025	-0.662	0.509

Dependent variable – 2007 TNCE
R^2 = 0.448, Durbin-Watson = 1.854, N = 440

Table 6 is the OLS regression for 2007 Study and Control 1 students. Study effects were significant. This is different from 2006. As is noted earlier, this was most likely due to there being no Type 2 Study students in 2006. This resulted in a very limited sample with which to work. The larger N as well as having both Type 1 and 2 for the study data may have resolved study effects for Study and Control 1 students. Again minority status was found to negatively impact scores in ways that are statistically significantly.

Table 7. OLS Regression for 2007 Study and Control 2 students.

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
2006 TNCE Score	0.677	0.020	0.752	33.385	0.000
Study Treatment	2.966	0.752	0.085	3.945	0.000
Minority	-2.292	0.834	-0.066	-2.747	0.006
SES	0.285	0.750	-0.009	0.380	0.704

Dependent variable – 2007 TNCE
R^2 = 0.555, Durbin-Watson = 2.015, N = 1142

Study effects were also significant for Control 2 and Study students (Table 7). This is a group that was not represented in 2006. Considering how much the Type 2 students in both groups were observed to differ in the descriptives, this is not surprising despite the Type 1 students appearing very similar in the descriptive statistics. This contrast has important implications for the findings from regression discontinuity analyses. For this OLS analysis, minority status is again significant but SES is not.

Table 8. OLS Regression for 2007 Study and Control 3 students.

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
2006 TNCE Score	0.724	0.020	0.740	36.203	0.000
Study Treatment	6.486	0.742	0.173	8.741	0.000
Minority	-3.168	0.732	-0.093	-4.329	0.000
SES	-1.064	0.682	-0.032	-1.560	0.119

Dependent variable – 2007 TNCE
R^2 = 0.538, Durbin-Watson = 2.047, N = 1421

Finally, the study effects are significant relative to the Control 3 students. The study students had almost 6.5 NCE points increase over non-study students. This effect is very consistent with the result from last year but in 2006 significance was not achieved due to low numbers in the study. Again, minority status and not SES is observed to have a significant and negative impact on scores.

Regression Discontinuity Analysis

The last analyses done for this report are based on regression discontinuity design. This analysis allows for examination of study effects by looking for a discontinuity at a cut-score (in this case, the TAKS passing score) when this score is transformed to have a zero value. The mathematical rationale is that under this transformation all other factors play no role at this cutoff point and thus any difference must be due to study treatment alone. All analyses performed included testing for higher order as well as interaction effects. Unless presented below, all such effects were insignificant and thus were not included in this report.

For regression discontinuity, transformed NCE scores continued to be used. For 2006, SES was chosen as the variable to control for student difference. While minority status has more often than not proven to have a significant negative effect, the homogeneity and even smaller N values of the samples used in regression discontinuity makes it difficult to control for minority status. While this is not ideal, it is forced by the limitation of the available data and could have the effect of under-reporting the effectiveness of the intervention. The N values were larger in 2007, however, and so minority status was used as a variable in all models tested.

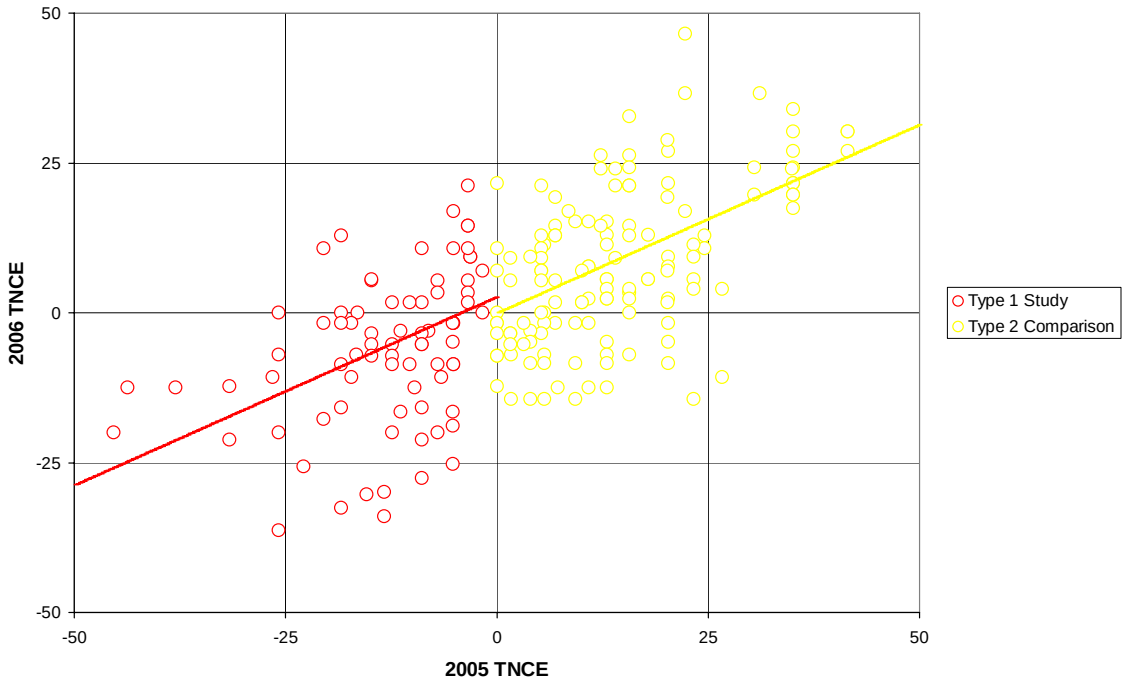
Table 9 is the regression discontinuity for 2006 Study and Comparison students. It is interesting to note that while the study treatment effect is ~3 NCE points and accounts for ~9% of the variance, it is still insignificant. This is most likely due to the low N values. Regression discontinuity requires a larger sample especially those close to the cutoff to be powerful. However, it is already known from the OLS that this effect should be significant. Since the OLS uses data from the whole sample, it is not as limited by the small N values as regression discontinuity. Following on these 2006 results, the 2007 data will be able to tell us more.

Table 9. Regression Discontinuity for 2006 Study and Comparison students.

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
2005 TNCE Score	0.629	0.070	0.730	8.994	0.000
Study Treatment	2.728	2.536	0.087	1.076	0.283
SES	-1.115	1.656	-0.35	-.673	0.501

Dependent variable – 2006 TNCE
R^2 = 0.443, Durbin-Watson = 1.313, N = 215

Graph 1. Regression Discontinuity for 2006 Study and Comparison groups



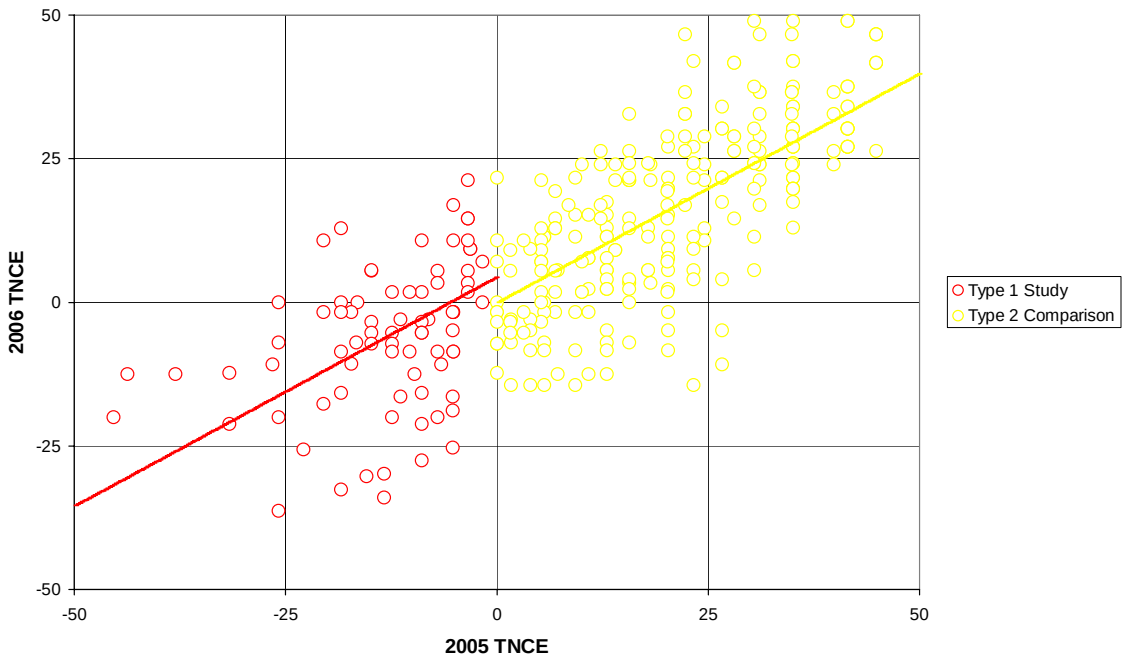
Returning to the 2006 results, the discontinuity at the cutoff is apparent in Graph 1. This discontinuity, however, is small and not significant. As explained earlier, this finding is different from last year's reported results because the NCE used in last year's report was generated using the frequency distribution of 2003. This was to create a baseline year for comparison. While this use of a baseline is often considered standard practice, given current test design (specifically the use of item response theory [IRT] for item development) it was decided this use of an arbitrary baseline would introduce distorting statistical artifacts into the analysis. This is seen most prominently in last year's report as a clustering of data points at the maximum NCE of 100. A baseline year was not used for this report and NCE scores were computed based on the frequency distributions for only the given administration of the test. In addition, the regression discontinuity last year used the pre-AP students in the Comparison group. For reasons outlined earlier, including this group of very distinctive students would introduce additional artifacts even if their inclusion would increase the overall N value (which was important in 2006). Just for comparison purposes, the pre-AP analysis using revised TNCE scores is included in Table 10 and Graph 2.

Table 10. Regression Discontinuity for 2006 Study and Comparison students including pre-AP

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
2005 TNCE Score	0.798	0.043	0.841	18.661	0.000
Study Treatment	4.500	2.136	0.091	2.107	0.036
SES	-5.098	1.293	-0.125	-3.942	0.000

Dependent variable – 2006 TNCE
R^2 = 0.688, Durbin-Watson = 1.187, N = 359

Graph 2. Regression Discontinuity for 2006 Study and Comparison groups with pre-AP students



As was seen in last year's report, this result is now significant. However, we no longer support their inclusion. Pre-AP students generally score much higher than non pre-AP students. This in turn changes the slope of the line to be much steeper such that when accounting for difference due to the study, the discontinuity would be increased. Thus, while removing the pre-AP students may tend toward under-reporting the effectiveness of the intervention, it is justified in terms of their very distinctive statistical attributes.

For the sake of completeness, the regression discontinuity for the Study and Control 1 was done even though it was insignificant on the OLS, shown in Table 11. It is not surprising then that results are not significant. It should be noted though that like the OLS, the result is trending towards significance.

Table 11. Regression Discontinuity for 2006 Study and Control 1 students.

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
2005 TNCE Score	0.653	0.084	0.745	7.763	0.000
Study Treatment	2.704	2.854	0.090	0.947	0.345
SES	-0.601	1.688	-0.020	-0.356	0.722

Dependent variable – 2006 TNCE
R^2 = 0.460, Durbin-Watson = 1.359, N = 183

Graph 3. Regression Discontinuity for 2006 Study and Control 1 groups

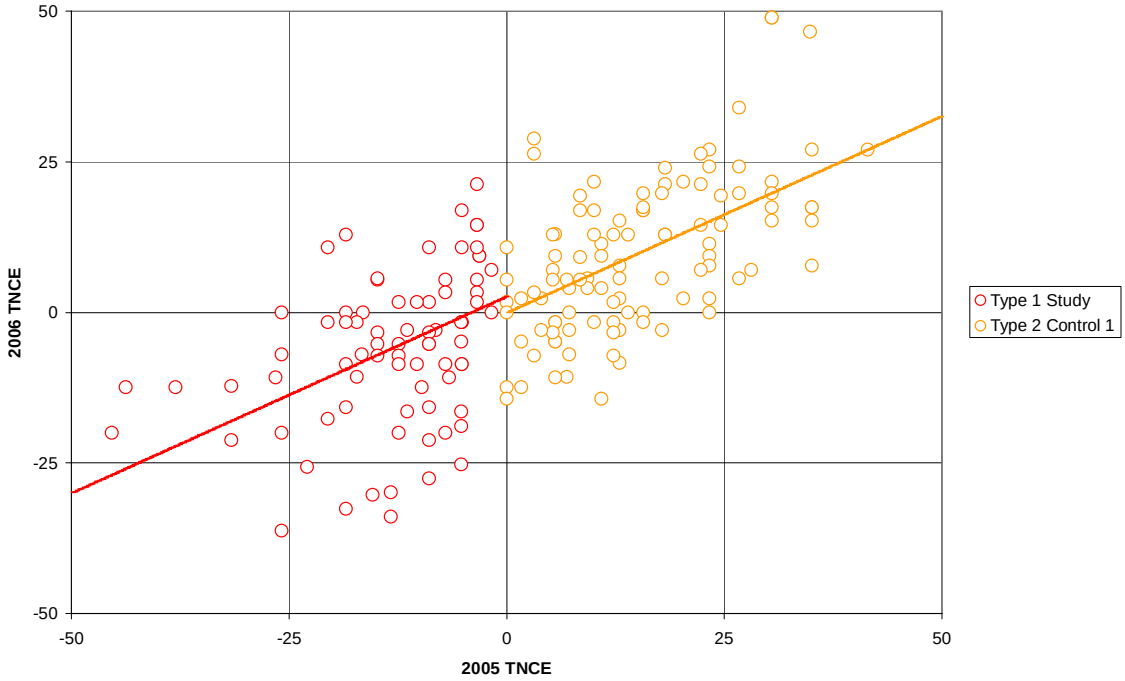


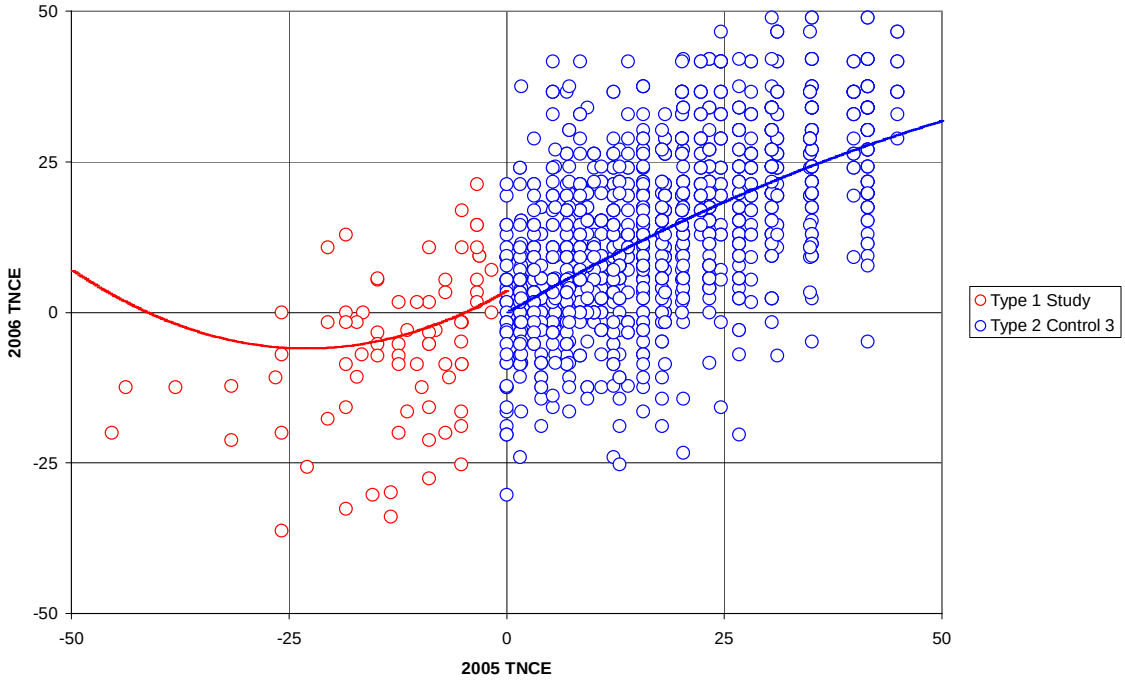
Table 12 is the regression discontinuity for the Study and Control 3 group, which again is insignificant as predicted by the OLS. Moreover, it is the only analysis to yield significant higher order and interaction effects.

Table 12. Regression Discontinuity for 2006 Study and Control 3 students

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
2005 TNCE Score	0.836	0.073	0.834	11.443	0.000
Study Treatment	3.659	3.609	0.054	1.014	0.311
SES	-3.011	0.652	-0.096	-4.619	0.000
Sq 2005 TNCE Score	-0.004	0.001	-0.193	-3.036	0.002
TNCE x Study Interaction	0.340	0.445	0.081	0.764	0.445
Sq Interaction	0.022	0.011	0.151	2.028	0.043

Dependent variable – 2006 TNCE
R^2 = 0.430, Durbin-Watson = 0.596, N = 1387

Graph 4. Regression Discontinuity for 2006 Study and Control 3 groups



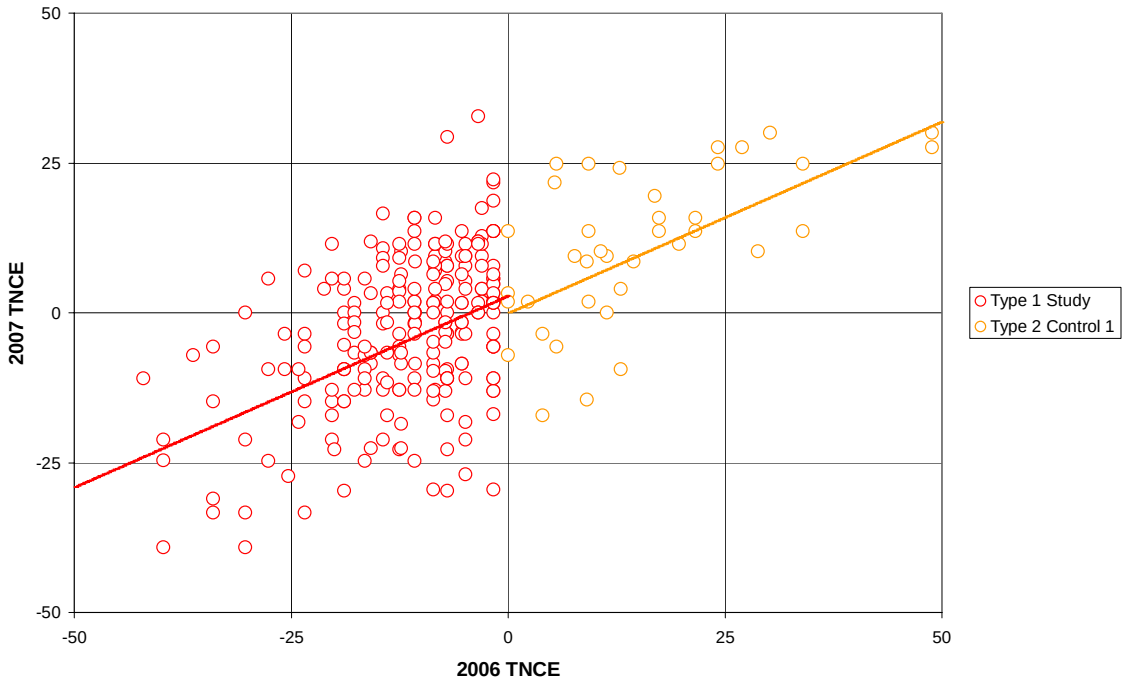
Next are the regression discontinuity results for the 2007 data. Table 13 shows that the study effect is not significant for Study and Control 1 students. This contradicts the OLS results where the study is indeed significant. While this particular analysis is trending towards significance (B = 2.594, explaining 7.6% of the variance), it is most likely due to low N values for the Type 2 Control 1 students. Graph 5 shows how few of those students there are from which the line is generated.

Table 13. Regression Discontinuity for 2007 Study and Control 1 students.

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
2005 TNCE Score	0.639	2.768	.623	9.000	0.000
Study Treatment	2.878	.071	0.074	1.073	0.284
Minority	-5.864	2.016	-0.142	-2.908	0.004

Dependent variable – 2007 TNCE
R^2 = 0.361, Durbin-Watson = 1.859, N = 274

Graph 5. Regression Discontinuity for 2007 Study and Control 1 groups



In addition, Type 2 Study students grew by over 5 NCE points whereas all other Type 2 students went down. Since OLS examines both Type 1 and Type 2 students for both Study and Control 1 groups, it is more able to make visible significant results, especially when combined with the fact that Type 1 Study students had the largest gains. To preserve the orthogonality needed for the regression discontinuity design, only the growth of Type 1 Study students are examined and compared to Type 2 Control 1 students.

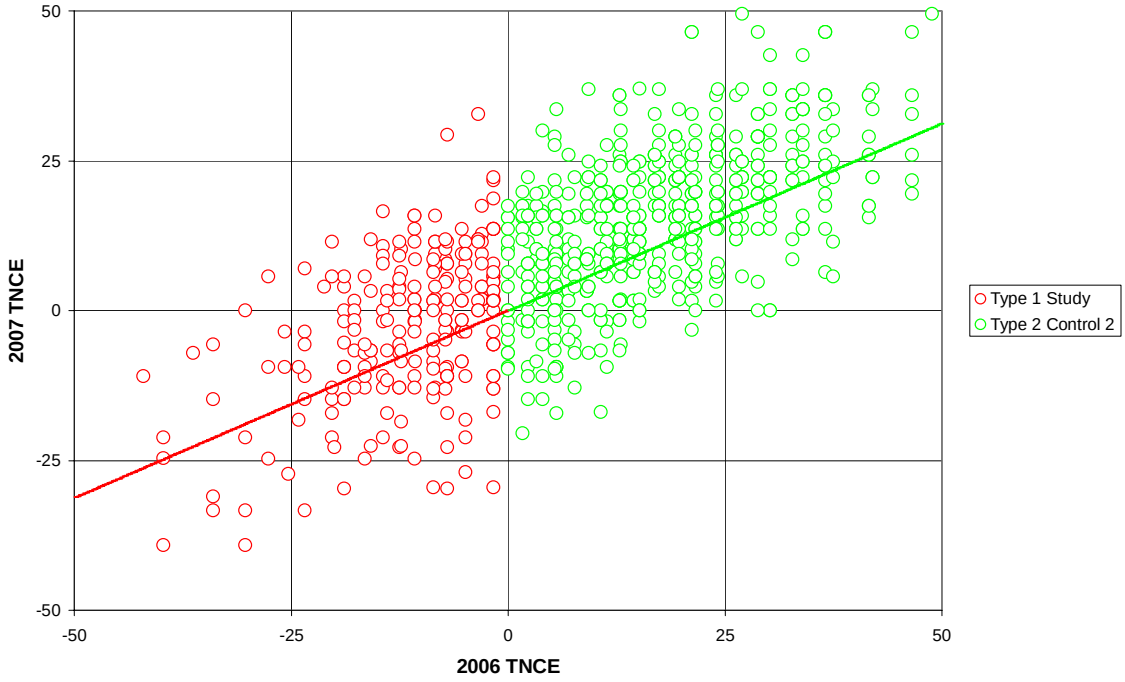
The analysis for Control 2 students gave curious results at first glance also. The OLS had indicated them to be significant. It should be recalled though that for Type 1 Study and Control 2 students, the change in NCE is much the same but that the Type 2 changes were drastically different with the Study going up by over 5 points and the Control 2 students going down over 2 points. It is probably for this reason that the OLS indicated significance. However, there is no significance in the regression continuity because as stated before, this type of analysis eliminates Type 1 Control 2 students and Type 2 Study students. This in effect cancels out the differences seen in the change from last year's NCE since the difference lies in the Type 2 students with the Type 1 students being almost the same for both groups.

Table 14. Regression Discontinuity 2007 Study and Control 2 students

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
2005 TNCE Score	0.626	0.032	0.724	19.534	0.000
Study Treatment	0.124	1.236	-0.004	0.101	0.920
Minority	-1.480	0.805	-0.047	-1.838	0.066

Dependent variable – 2007 TNCE
R^2 = 0.547, Durbin-Watson = 2.074, N = 820

Graph 6. Regression Discontinuity for 2007 Study and Control 2 groups



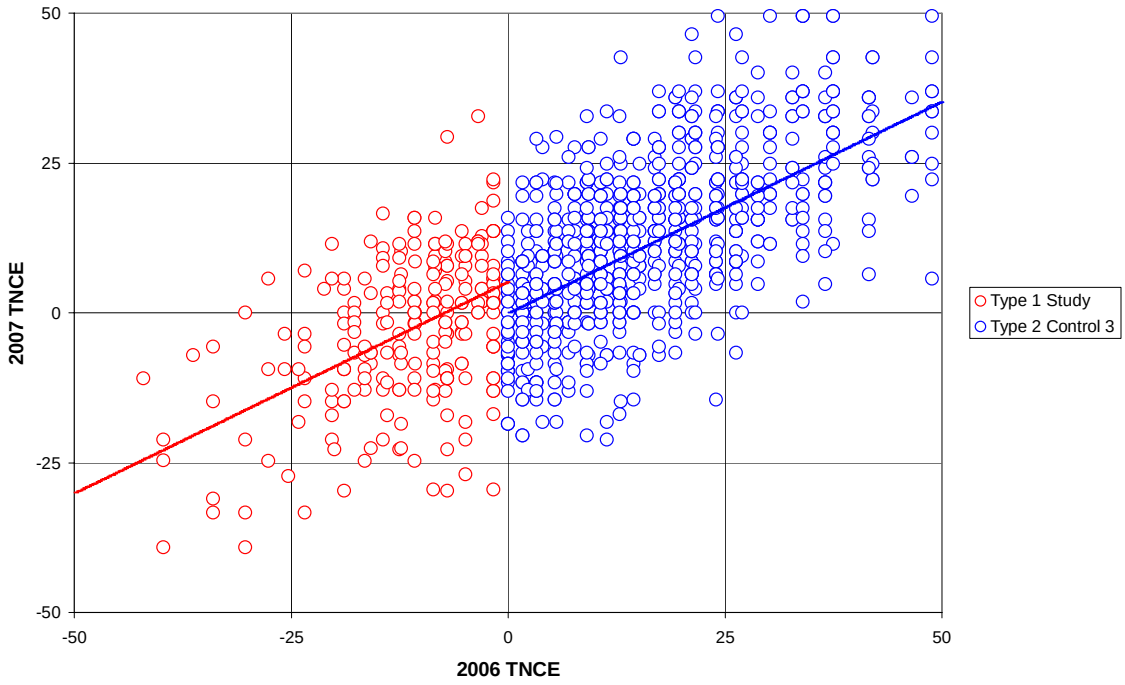
The last analysis is a regression discontinuity with the 2007 Study and Control 3 students. Inspection of the descriptive statistics indicates significant differences in outcomes for students. Correspondingly, the results for the regression discontinuity analysis is significant.

Table 16. Regression Discontinuity 2007 Study and Control 3 students

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
2005 TNCE Score	0.706	0.030	0.759	23.560	0.000
Study Treatment	5.263	1.166	0.143	4.512	0.000
Minority	-3.450	0.744	-0.110	-4.639	0.000

Dependent variable – 2007 TNCE
R^2 = 0.504, Durbin-Watson = 1.970, N = 1030

Graph 7. Regression Discontinuity for 2007 Study and Control 3 groups



Conclusions

In conclusion, under OLS analyses the study intervention is effective in raising both Type 1 (students who failed the previous year TAKS) and Type 2 (students who passed the previous year TAKS) students' mean NCE scores. This lends significant support for the versatility and inclusiveness of the intervention when it comes to classroom use. Due to this increasing of the Type 2 Study students' scores and lack of growth in all other Type 2 students, OLS regression analysis always yield significant results, but regression discontinuity often did not. The closer the Type 1 Control students were to the Type 1 Study students, the more likely the regression discontinuity would fail to find significance.

Future work, to validate some of the implications of these analyses, should examine what is happening in the Control classes. This is especially true for the Control 2 classes which in this analysis resemble the Study classes the most at the Type 1 level.

Regression discontinuity analyses did show significance at the district level comparison. In general across OLS and RDD analyses, when significance was found the effect of the intervention was in the four to six point range for improved NCE score on a 100 point scale. Even when significance was not reached, the results often were trending in this range. This convergence of results across complementary methodologies lends further credibility both these findings and to the methodologies developed for these analyses.

To conclude, the overall results indicate that the *MathForward* intervention resulted in scores of students below passing in one year improve their scores by 4-6.5 points in the subsequent year.

In contrast to other forms of intervention that result in some improvement in outcome for underperforming students but at the apparent expense of students scoring above the passing level, the results of this study suggest scores for all students in classes using the *MathForward* program improved. All students appeared to benefit from participation in the *MathForward* program.

References

Cook, T. D. & Shadish, W. R. (1994). Social Experiments: Some developments over the past fifteen years. Annual Review of Psychology, *45*: 545-580.

Shadish, W.R., Cook, T. D., & Campbell, D.T. (2001). Experimental and Quasi-Experimental Designs for Generalized Causal Inferences. Houghton Mifflin Company, Boston.

Texas Education Agency (2007).
http://www.tea.state.tx.us/student.assessment/reporting/freq/

Trochim, William (2006) http://www.socialresearchmethods.net/kb/quasird.htm