DOES IT PAY TO GET AN A?
SCHOOL DISTRICT RESOURCE ALLOCATIONS IN RESPONSE TO ACCOUNTABILITY RATINGS

Steven G. Craig

Scott A. Imberman

and

Adam Perdue

Department of Economics
University of Houston
Houston, TX 77204-5019

scraig@uh.edu
simberman@uh.edu

Preliminary and Incomplete
Please do not cite without permission.

Abstract

This paper examines a hitherto unexplored, but central, part of the impact of the new school accountability standards, which is the response of school districts to accountability ratings. We carefully model schools on the margin of the well-defined ratings cutoffs, which allows us to specify a regression-discontinuity design that estimates the impact of school ratings on the allocation across schools of financial and personnel resources by school districts. Preliminary results suggest that school districts provide scant rewards for the performance of the schools' students on state-wide standardized exams.

November, 2008

I. Introduction

This paper studies how school districts allocate resources across schools in response to the accountability rating of schools. Answering this question is the necessary condition for understanding how schools and districts respond to accountability standards. That is, school administrators are much more likely to respond to accountability ratings if they believe there are consequences for their schools. So for example papers by Figlio and Winicki (2005), Jacob (2005), Cullen and Reback (2006), Figlio (2006) which find that schools change their internal behavior in response to accountability standards may not make sense, if there are no external incentives to motivate internal school activities. While one incentive may be a potential loss of enrollment and the associated funds school districts may also create their own incentives by providing rewards via additional resources for schools that attain high ratings. On the other hand, school districts may try to prop up schools with lower ratings by increasing their resources at the expense of the higher rated schools. Our paper identifies the extent to which either of these potential responses to accountability ratings occurs.

We analyze the school district allocation decision by employing a detailed data set on Texas schools over the years 1997-2007. The accountability standard in Texas during this time changed in 2002. The results in this preliminary paper therefore cover only the initial standard, called TAAS (Texas Assessment of Academic Skills) for the five year period 1997-2002. The final version will include results from the ostensibly more rigorous TAKS (Texas Assessment of Knowledge and Skills) period after 2002 and will thus provide comparisons between the two systems.

Accountability standards in Texas consist of firm lines based on the pass rates of students

on a standardized exam administered by the state and dropout rates. There are four rankings, exemplary (E), recognized (R), acceptable (A), and low performing (L).[1] Higher scores by one student do not compensate for lower scores by another; the criterion is based solely on whether a student passes the exam.[2] The available school-level financial data includes information on expenditure on a variety of categories, including instruction, counseling, extra-curricular, etc., as well as on the quantities of particular types of employees. Our objective in this research is to see how the allocation of the fixed district resources is affected by the ranking achieved by each school.

There are three general theories (or strategies) about how school districts might re-allocate resources to schools in response to the accountability standards, the "push-up" theory, the "reward" theory, and the "school specific" theory. In the "push-up" theory, districts would award schools that are below the standard with extra resources so that they are able to reach the minimum standard. A disadvantage of this strategy is that schools which are able to reach the standard would then see a reduction in available resources. An alternative theory is that schools which are able to get over the barrier to the next ranking would be "rewarded" with additional resources for the school. The disadvantage of this theory is that schools which are below the barrier would have fewer resources to attain the objective. The final possibility is that school districts believe that attaining the accountability objectives is a school specific problem, and that no resource reallocations would result. This behavior might be consistent with the strong

[1] The "Low Performing" ranking under TAAS was renamed to "Academically Unacceptable" under TAKS.

[2] That is, there are no standards for average score levels, although recently students are being recognized for "commended" performance.

Rawlsian criteria inherent in the accountability standard, which is based only on passing rates. Our research objective is unique, in that we attempt to differentiate between these potential response functions by examining a broad array of school specific inputs. While Rouse, Hannaway, Goldhaber, and Figlio (2007) provide evidence from a unique survey of schools that identify a number of policy changes that are induced by low rankings, there is very little evidence on how schools and districts re-allocate resources when they get lower rankings. Jacob (2003) does look at how school resources in Chicago adjust to the imposition of an accountability system and finds evidence of a shift in expenditures to non-ancillary instruction amongst schools with low pre-accountability test scores but overall there was little change. He does not, however, look at resource allocation responses to the assignment of rankings.

Inputs we examine in the data include expenditure by function and by program, as well as numbers of employees. Some of the functional expenditures include instructional, teacher development, leadership, and two types of counseling. The programs include career and technology, in addition to basic instruction. The employee counts are for teachers, teacher aides, and administrators. We normalize all of these inputs by enrollment.

Our empirical strategy is to use a regression discontinuity research design, which identifies the effect of being on one or the other side of the accountability ranking. Ultimately these estimates will be compared to a standard regression model. We separately analyze each of the three margins, the LA border (low performing to acceptable), the AR border (acceptable to recognized), and the RE border (recognized to exemplary).

A key data input is therefore to determine whether a school is at the margin between two rankings. We determine the distance from the ranking boundary by counting the number of

students that would have to pass the exam, but did not, for a school to rise one rank. Similarly, we count the number of students that if they failed, would lower the rank of the school. These two measures, the distance in terms of number of students above the border, or the distance below, are not symmetrical, because of the Rawlsian evaluation inherent in the Texas accountability standards. Specifically, a school's overall ranking is the minimum evaluation based on a number of criteria. That is, the percentage of students that pass the standardized exam for each group by race and ethnicity, economically disadvantaged, and overall, for each subject (depending on grade level), is calculated. Each cell in this matrix, except overall, must have either thirty students and at least 10% of alls students tested or fifty students and less than 10% of tested students to count towards a school's ranking. Of the groups that qualify, then the lowest percentage passing for each student type and subject combination determines the school's overall ranking. Thus to fall to the next lower ranking, a school could have as few as one additional student in a specific race-subject combination fail. On the other hand, to rise in the ranking, all of the cells with passage rates below the next border would have to be raised. We build a model to estimate the number of students required to change each school's ranking from above and below, while accounting for the exceptions to the general standards.3

Our empirical strategy is to estimate our model for all of the data in Texas with the above criteria. It is possible, however, that the objective function of the school districts are not identical across the state, in which case the pooling assumption would be violated. We attempt to test this assumption by estimating our model for the largest school district in the state, and for the aggregation of the five largest. Because the regression discontinuity design requires a

[3] The exceptions include attendance, drop-outs, and improvement over time.

significant number of schools at the ratings borders, these tests are not decisive. Nonetheless, they suggest that the response of school districts is not widely disparate across the state, and that our pooled Texas regression results are likely to isolate how typical school districts are re-allocating their resources to create incentives for schools to do well relative to the accountability standards created by the state government.

Our results show that school district governments are creating very few incentives for schools in Texas to meet the accountability standards set by the state TAAS system, and least through the resource allocations between schools. The exception at the elementary level is that extra counseling resources appear to be directed at the Low performing schools. On the other hand, the few small incentives we observe where schools are rewarded for being above a rating boundary are directed as much at the administration of the school as at the learning resources for children. The specific categories that are estimated to increase vary considerably, but the general finding holds for all three borders in our sample, and for all three school types. These early results therefore are consistent with school districts setting small incentives to encourage schools to improve their test scores by rewarding them with resources due to good results.

Our paper proceeds first by discussing the prior literature. Then we explain the regression discontinuity research design. A necessary input is to determine whether schools are on the margin of one rating level or the other, and we discuss our process for doing so. A data section briefly describes the Texas data. Then, we briefly examine an OLS specification with dummy intervention for rating, which clearly shows the need to account for simultaneity. Next,

we provide our results from our regression discontinuity analysis. A final section summarizes and concludes.

II. Empirical Methodology

The objective of our research is to discern the underlying objective function of the school districts in Texas when faced with accountability pressures. To do this, our model will ascertain the manner by which districts respond to the schools' success in achieving high ratings. Our empirical model will show whether districts create incentives by rewarding schools which perform well, or instead whether districts respond by putting extra resources into low performing schools.4 Irrespective of this decision, however, the method by which school districts create incentives for schools is essential for understanding whether, and how, individual schools change their behavior in response to accountability programs. Generally, researchers have found a clear improvement in test-score performance when states implement accountability systems both on high and low-stakes exams (Hanushek and Raymond, 2004; Hanushek and Raymond, 2005; Jacob, 2005; Carnoy and Loeb, 2003), and the impacts are particularly strong when accountability regimes include the threat of giving vouchers to students at low-performing schools (Chakrabarti, 2008; Figlio and Rouse, 2006). Nonetheless, these gains do not necessarily accrue to students at all points on the test-score distribution (Reback, 2008; Neal and Schanzenbach, 2007).

[4] We leave for later attempting to discern whether districts desire to have the largest number of schools with high ratings, and thus focus on small schools, or alternatively, maximize the number of students experiencing education in a highly rated school.

An important related question is whether the test score gains represent actual improvements in learning, or are simply a result of teaching test taking strategies. For example, there is substantial evidence of schools trying to prevent low-scoring students from taking exams by reclassifying them into exempt categories or instituting longer suspensions during testing windows (Cullen and Reback, 2006; Figlio, 2006; Jacob, 2005). Figlio and Winicki (2005) even find that schools increase the caloric intake of their students on exam days to improve scores. The extent to which these type of responses might be expected to be representative is in part dependent on the importance that school districts impart to the accountability outcomes.

To understand school district resource allocations it is essential to determine the margins along which they are operating. To do this, we assume that the distance from a margin is based on the number of students in each group and subject that need to switch from fail to pass or vice-versa to reach the margin.[5] Specifically, we create two variables, N_a and N_b, which describe the distance to the rating border above, or below, a school's current rating.[6] Under the TAAS system ratings could be E (exemplary), R (recognized), A (acceptable), and L (low performing).

[5] There are two implicit assumptions we are making here. First, is that the costs of adjustment do not depend on school size. Second, the cost of moving each student-group-subject unit from fail to pass is constant. For example, suppose a school is below a threshold for white and economically disadvantaged math and reading. Then a student who is both white and economically disadvantaged and fails both math an reading will cost four times as much to improve as a student who is white but not disadvantaged and only fails math.

[6] In some cases a school will have N_a and N_b indicating an L rating even though the school is actually A. This is because some schools are able to get to an A rating via a "acceptable improvement" clause that allows the school to use increases in passing rates rather than levels to get an A rating. While school districts may only care about this lower threshold, by 1999 each school was required to reach the level thresholds and this was known to schools in 1997 and 1998. Thus, since schools ultimately had to reach the threshold, we assume they cared about the level rather than the acceptable improvement. Later versions of this paper will incorporate acceptable improvement into specification checks.

We define N_a as the number of students above the minimum for the school's current rating, so if $N_a + 1$ students switch from passing to failing the state exam, the rating of the school would be lowered. Similarly, N_b is the number of students by which a school is below the border to the next higher rating, so if N_b number of students switch from failing to passing the state exam then the rating of the school would be raised.[7]

Because the rating criteria are based on the minimum performance of the defined student-subject groups, however, the two measures are not symmetric. If the performance of any group falls below the school's current rating, the school would receive the lower rank. Thus N_a is based on the how many students would need to switch passing status in the subject-group that is closest to the border of the next lower ranking. On the other hand, to raise a school's ranking to the next level, the performance of all of the subject-groups that count towards the rating would have to improve.

There are a number of significant features of the Texas TAAS system, however, that must be understood before the variables N_a and N_b can be measured. Table 1 describes the requirements for the sub-populations and subjects for each type of school, elementary, middle, or high school. Students are defined based on their race and ethnicity into white, black, and Hispanic. In addition, students are identified as economically disadvantaged if they are on free or reduced price lunch (and/or breakfast). Finally, the performance of all the students in the entire school is counted as a group. Other races are not evaluated as groups, though the students are still included in the "all-students" group. For each of the student groups, passing rates are

[7] Dropouts are also included as a measure for middle and high schools, in which case N_a and N_b become the number of additional dropouts needed to drop the rating plus one and the number of dropouts need to return in order to increase the rating, respectively.

divided by subject area. Reading, math, and writing exams are given in all years while in 2002 social studies is added for 8th grade students. Middle and high schools are also evaluated on dropouts. Thus a school's rating is based on the minimum student pass rate in each of the sub-populations for each subject. If the subject-group contains less than 30 students or between 30 and 49 students but less than 10% of the all-student count for that subject then it will not count for the rating. In addition to the basic rating described above, there are exceptions to a school's rating based on attendance, drop-out rates, improvements over time, and an appeals process.[8]

Our regression discontinuity design exploits N_a and N_b to identify whether the district changes the resource allocation to schools based on which side of the accountability ratings line a school falls.[9] Ignoring the exceptions for attendance and drop-outs, and assuming that the required improvement and appeals process is non-binding, the school accountability ranking formula from Texas is:

$$SR = \min (SR_i) \tag{1}$$

where SR_i is the school ranking for each student group - subject combination.

[8] From 1997-98 through 1999-00 there was also an attendance requirement. Schools that failed the requirement were given an A rating if they were otherwise supposed to get an R or E but could not be given an L due to attendance. Schools could also appeal to use their current year attendance rates instead of previous year if that improves their ranking. Since this requirement is based on average daily attendance and not student counts, it is impossible to combine this measure with the other subjects in our current framework. Thus we do not include attendance in our distance from the margin calculations. In some cases a group is exempted from the dropout requirement if they make substantial progress. We do not, at this time, include this waiver in our model. Additional appeals are mostly used to correct data mistakes.

[9] At this point, we model the treatment as sharp, and not as fuzzy, see Hahn, et. al. (2001).

To calculate the margin by which a school is over or under the line requires that the ranking of each student-subject combination be considered. We define N_a^i as the number of students that passed the exam in group-subject combination i above the minimum number needed to attain the school's overall SR. To determine the margin by which a school just achieved its ranking, we measure the minimum number of students who would have to fail the exam, rather than pass. This margin depends only on the lowest performing group-subject combination, since if even one combination falls below the line illustrated in Table 1 the entire school's ranking will be lowered. Thus we measure N_a as:

$$N_a = \min (N_i) \qquad \forall \quad SR^i = SR \qquad (2)$$

To determine the margin by which the next ranking could be attained, define N_b^j as the number of students who would have to pass the exam, but did not, to achieve the next higher SR in group-subject combination j. Because of equation (1), the minimum number of students whose exam performance would have to change is the number that would raise the school ranking of every group-subject combination to be the next higher school rank. Thus:

$$N_b = \sum_j N_b^j \qquad \forall \; SR^i = SR \qquad (3)$$

That is, we count the number of students in each group-subject combination where the school ranking is just equal to the school's overall ranking. If the ranking of a group-subject combination is already above the school's overall rank, it will not contribute to the calculation of N_b, because of the maximin criterion in (1).

Thus our empirical specification is:

$$R_i = \beta_1 f(N_a) + \beta_2 f(N_b) + \beta_3 SR + \varepsilon \quad (4)$$

where R is the school specific resource (either dollars per student or inputs per student), and N_a and N_b are the number of students by which the school is above (N_a) or below (N_b) the rating border as defined by (2) and (3) above. We allow in our regression higher order polynomials of N_a and N_b for generality (indicated by the functional). SR is a dummy variable which describes whether a school is above, or below, a given rating. To explore whether districts have different implicit weights on different boundaries, we run our regression (4) separately for each boundary, that between L and A schools (the LA boundary), between A and R schools (the AR boundary), and between R and E schools (the RE boundary).

The final attribute needed to estimate (4) is to determine whether a school is at the margin of achieving a particular ranking. The trade-off is that wider margins are both less likely to be randomly assigned between rankings, and more difficult to specify. A narrower definition of the margin, however, gives fewer observations of schools at the margin. With that in mind, we present results below that consider 20 and 30 students above or below the margin.

III. Data

Our data covers all the schools in the state of Texas and comes from three public datasets provided by the Texas Education Agency. The first is the Academic Excellence Indicator System (AEIS) which provides data on school staffing, enrollment, and student demographics. The

second is the Public Education Information Management System (PEIMS) financial reports which provide expenditure data. The third is the TEA accountability rating reports which provides the data that is used to calculate accountability ratings. In total, the data covers 7,053 schools in 1997 and 7,956 schools in 2005, which combine the TAAS accountability system (1997-98 through 2001-02) and the TAKS system (2003-04 through 2005-06). There were 59 schools with an L rating in 1997-98, and 267 in 2005-06, although the number did not increase in all years. In this version of the paper we focus on the TAAS period from 1997-98 through 2001-02. Later versions will also include an analysis of TAKS and potentially earlier years under the TAAS system.

Because we are interested in school district allocation choices between schools, we restrict our sample to districts with more than one school of each school type, elementary, middle, and high school. For the core of our analysis, we delete the alternative schools, which generally are targeted to specific groups of students, and which operate under a separate accountability system. We also restrict our sample to schools with at least 200 students, and to districts with more than one school of each type. This leaves us with 13,821 observations on 3048 elementary schools, 4,157 observations on 952 middle schools, and 1,464 observations on 341 high schools.[10]

The resource data has three sets of information. From PEIMS we get school-level expenditures classified by functions, and programs. The functions we examine are instruction, instructional resources, curriculum and staff development, instructional leadership, school

[10] After the sample restrictions there remain 43 observations on 23 schools that do not fit into the elementary, middle, or high grade category that we also drop.

leadership, guidance and counseling, social work services, and extra-curricular activities. The programs we examine are basic instruction, and career and technology.11 The AEIS data provides full-time equivalent staffing counts broken into teachers, aides, and administration. Table 2 presents the means of the data. Table 2A presents the transition matrix between different rankings, the high degree of movement suggests that the pooled analysis we perform here, where each year is treated separately, is appropriate.

IV. Empirical Results

We present two tests for the impact of the TAAS accountability system on the allocation decisions of the school districts in Texas. In Table 3, we show the OLS estimates from a linear regression of expenditures per student in a variety of categories, with each school's ranking indicated by a dummy variable. There are several reasons to believe the OLS estimates are biased, and are not useful for reaching any conclusions. For example, school districts may provide more funding to low-performing schools in general, regardless of rankings. On the other hand, high quality schools may already be using high levels of expenditures to achieve their current quality level. The OLS regression results are consistent with the first theory. Higher ranked schools are found to have lower expenditures per student over a broad range of categories, as well as lower staff-student ratios.

Thus, in Tables 4 and 5 we now turn to our regression-discontinuity results which provide causal estimates of the impact of accountability ratings on resource allocations. The key

[11] We omit the gifted and talented program, because of concerns over the endogeneity of the number of students (from parental selection of schools).

results in this table are those on the intercept terms, which illustrate the financial allocations to each of two otherwise identical schools, where one is marginally above and the other is marginally below the rating cutoff.

A key element for interpreting the regression discontinuity intercept coefficients is the definition of which schools are on the boundary between ratings. We explore a wide variety of definitions for which schools are on the margin. The tables present the estimates assuming the margin is defined by either 20 or 30 students on either side of the rating boundary.[12] We find our results our relatively robust to other boundary definitions, we explore between 10 and 50 students. The other element of sensitivity analysis is the degree of the polynomial used to describe the distance to the boundary. The results in the tables use a third degree, but are qualitatively robust for up to a 6 degree polynomial. Finally, we find it is useful to estimate our model separately for elementary schools, while we pool the middle and high schools. The pooling is necessary because we present our results separately for each boundary; the LA border, the AR border, and the RE border.

The objective of the estimation is to determine how school districts systematically allocate resources across schools in response to their accountability rating. The advantage of the detailed expenditure and input data is that we are able to present a somewhat nuanced story, because there are distinct differences over incentives for achievement, as well as assistance to reform poor performance. These distinctions vary significantly depending on the type of school,

[12] We also explore the asymmetry between N_a and N_b, because as described in equation (3) N_a is defined only for the most marginal group-subject combination, while N_b aggregates all of the students that are below a boundary as in equation (2). Thus we tried specifications that allowed the N_b distance to be between two and four times as large as the N_a distance. The

and even more on which of the ratings boundary is being examined. On average, however, we find that school districts are creating small incentives for schools to respond to the state administered school accountability standards.

1. Elementary Schools

Elementary schools on the border between low performing and acceptable would be expected to receive significant attention, especially given the recent emphasis on early childhood learning (such as Head Start). We find that schools on the L side of this border receive statistically significantly more money per student for guidance and counseling than do otherwise identical schools, although there is no evidence that total expenditure per student is higher. The point estimate of -$49.30 is 30.8% of the average guidance and counseling budget, which indicates the L rated schools are receiving significant attention. The extra funds are despite that there is no statistically significant increase in total or instructional expenditure for the L schools.[13] For schools that are just able to achieve an acceptable rating, however, we find that extra-curricular funding is higher. While the coefficient is small compared to the guidance and counseling coefficient, the positive number is 45% of the average budget. Interestingly, however, no other school input is found to be re-allocated on this border.

---

regression results appear very robust to large variations in how the border margins are defined, so we leave the definitions symmetrical.

[13] The negative coefficient for instruction is not robust to the definition of the margin. The coefficient is positive when the margin is defined as 10 students, is significantly negative when defined as 20, and is a small positive number when defined as 50. In contrast, the guidance and counseling result is negative for all definitions, and only for 10 and 50 does it lose statistical significance.

Figure 1 presents the marginal impact on the guidance and counseling budget on the LA border for elementary schools. This figure uses the 30 student definition of the boundary. There is clearly much more noise on the low side of this border, our strategy is to attempt to add two further years of data farther back in time to attempt to get more schools rated L, and to refine whether the intercept statistical result is the correct interpretation.

On the AR border, we find that schools which are just able to achieve the higher rating receive extra resources in curriculum and staff development, school leadership, and guidance and counseling. These results are portrayed graphically in figures 2 - 4. The total of $35.60 per student is about 2% of the non-instructional budget. For schools of average size (517 students), the total amount is not large by any means. It is possible to view these resources as benefitting the employees of the school, as for example staff development funds may increase the market value of the teachers that are trained. Similarly, extra funds for leadership may be more of the nature of administrative awards rather than directly to the benefit of students.

The reward for reaching the top ranking is considerably higher (see figures 5 and 6), because in addition to the non-instructional budget increase that is comparable to that for the AR border ($31.70 vs. $35.60); the schools that receive the E rating receive a substantial boost to their instructional budget (figure 7). Schools that receive the top rating are found to receive $72.90 more in instructional expenditures, about 2% of the total. The programmatic expenditure row shows an even larger increase in basic expenditures, $107.60 (figure 8). For a school with the average number of students, this amounts to enough funding to hire about one more person. What is interesting is that our results show that on average this extra person is an educational aide or an administrator, rather than a teacher.

2. Middle and High Schools

Table 5 provides results for middle and high schools pooled due to the much smaller number of schools.[14] For the LA border, we need to be cautious when interpreting results as, despite pooling, there are very few of these schools that get rated L, hence almost all of the observations are above the border. With that caveat in mind, we nonetheless find some evidence of resource re-allocation on this border. The largest impact is on spending for career and technical programs, which is higher for A than L schools. The coefficient of $172 represents over 82% of the mean expenditures per student. This result suggests that low performing schools have their technical training quite reduced, while schools that perform at the acceptable level receive a substantial increase. This response is particularly interesting in that it conforms well to the intuition that since L schools would need to focus on test-preparation, they would shift students' time from non-test related activities. This conclusion needs to be cautiously interpreted, however, because as shown in Figure 9 there is a lot of dispersion in among schools rated L. Consistent with the story from the statistical results, however, we also see an increase in teacher-student ratios for A schools which could correspond to increased hiring of career/tech teachers as the amount of the increase in that funding is enough to hire two teachers for the average sized school (677 students). For the AR and RE borders there are no statistically significant results. Thus, it appears that, with the exception of career and technical program

[14] Un-pooled results were generally similar with the exception of guidance and counseling estimates which were negative for middle and positive for high school. High schools also showed some weak evidence of increases in instruction and instructional leadership for R schools relative to A schools.

spending, the entirety of the district level response to accountability ratings is at the elementary school level.

3. Large Urban Districts

One concern with pooling our analysis over the entire state is that the school districts have different preferences, or strategies, with responding to the school accountability system. By this concern, the reason for not finding strong re-allocations is that some districts may be concentrating resources on the low performing schools, while others are directing resources to the successful schools. To test this idea, we examine the behavior of the five largest school districts which are all with the major metropolitan areas of Texas, Houston, Dallas, Fort Worth, San Antonio, and Austin. While this analysis more accurately reflects the possibility that the objective functions of school districts differ, it has the disadvantage of many fewer observations, especially at the margin of our regression discontinuity design.[15]

Tables 6 and 7 present the results from this analysis. While these results combine five separate districts, the large urban districts are likely to have more in common than the many more rural or smaller town districts from the rest of Texas.[16] Nonetheless, we only have enough observations on both sides of the border to look at elementary schools, and perhaps the AR border in high school.

[15] We also conducted an analysis of the largest school district in Texas, Houston Independent School District (HISD) on its own. These results, found in Appendix Table 1, show large re-allocations, but the bottom of the table shows that there are very few observations on one side of each border. For this reason, we view these results as statistically unreliable.

In spite of these limitations, the results are generally supportive that the objective functions of school districts are generally similar, although clearly also they are not identical. For elementary schools, we see results in Table 6 on the low-acceptable (LA) and acceptable-recognized borders that are similar to the overall state results. Like the overall state results, we see that guidance and counseling expenditures are (weakly) increased for the low performing schools. The 'reward' portion comes not, as in the overall state results, for extra-curricular activities, but instead for curriculum and staff development, and a small amount for social work.[17]

The AR border results show, if anything, a stronger incentive response to schools that are able to get on the higher side of the border. The overall response is estimated to be $594 per students, of which less than half ($218) is estimated to go for instruction. These amounts, however, are much larger than the individual categories, although consistent in sign. We also observe that the number of teachers increases, and weakly that the number of administrators increases. For the RE results, on the other hand, we do not find any impacts that are statistically significant at the 5% level, which differs from the state-wide results which showed small incentives for rating improvements. One possible explanation for these results is that large urban districts have more low and acceptable schools relative to the other districts - 70% of school-years we study in the five largest urban districts have an L or A ranking compared to just 42% in other districts - and thus they care more about the getting schools over lower boundaries than the

[16] Over the time-span of our data there were a total of 1063 school districts in Texas of which 376 have multiple elementary schools, 226 have multiple middle schools, and 82 have multiple high schools.

[17] The impacts on social work may be like the administrative increase we see elsewhere for high schools, in that the administrative support in the school is increased.

high R/E boundary. We thus conclude from these results that for elementary schools, the large urban districts are generally similar to the overall state in that incentive effects but they focus more on improving the lower-ranked schools. For middle and high schools in the largest five urban districts we find little evidence of re-allocations in response to accountability ratings which is also consistent with what we find in our state-wide analysis.

V. Summary and Conclusion

Our objective in this research has been to establish how school districts respond to accountability ratings. Specifically, we investigate whether school districts offer their schools incentives in response to student performance on the factors that determine rankings. The incentives we examine here are resources to the school, which can include both instructional and other resources. Whether school-wide resources are a sufficient inducement to educators to change their behavior is an open question which depends on the career paths of both administrators and teachers, and in a much more complicated way on the response of parents and the school financing consequences.

Our research design and highly detailed data on school's financial records offer a unique opportunity to find out whether school districts provide incentives to schools. We do this through a regression discontinuity design, which compares schools just barely above and below the cutoffs for the four school rankings. This strategy ensures that any bias due to the endogenous relationship between school performance and funding is negligible. To implement this design, we pool all of the school districts in the state with multiple schools at each level (elementary, middle, and high school). Nonetheless, we present both OLS results on the whole

sample and estimates that restrict our sample to the largest five urban districts. The similarity of the state-wide and five-largest districts estimates suggest that pooling across the state generates little aggregation bias. As such, our results rather indicate that school districts are only offering small financial and resource incentives to improve accountability rankings.

We find that on the low/acceptable border, school districts appear to be directing counseling resources towards the low performing schools at the elementary level. To be consistent with maximizing behavior, this action suggests that central school administrators believe test result shortcomings are a result of student behavior, not because of the level of educational resources. In middle and high schools we do see a drop in career and technical education spending in schools that receive a low rating which is consistent with districts and schools refocusing these students' time towards test preparation.

On the acceptable/recognized and recognized/exemplary borders for elementary schools, we do find increases in some categories of expenditures when schools receive higher ratings. Since these do not appear to be offset by drops in spending in other categories, this is consistent with districts providing incentives for higher ratings. Nonetheless these increases are small relative to the schools' total budgets and thus it is unclear how likely these incentives are to induce behavioral changes in the schools. Results for middle and high schools, provide little evidence that school districts provide incentives or punishments for schools on these higher rank borders.

An important caveat is that we do not examine the behavior within the school. Nonetheless, the general thrust of these results suggest that central school administrators believe

student performance are outside of their control, as resource allocations between schools do not appear to be generally motivated by the accountability ratings.

A final possibility which we hope to incorporate is that the state school finance system has diluted the incentives that school district administrators themselves would otherwise face. That is, presumably one outcome from high performing schools is greater parental demand. If such demand results in higher property values, as shown to be the case in Florida (Figlio and Lucas, 2004) though not in Charlotte, NC (Kain, Staiger and Samms, 2003), then for school districts which rely on property taxes, there is a potential increase in tax revenue from higher ratings. However, in many districts in Texas, school finance equalization eliminates any revenue gain from higher property values due to the "recapture" system of school finance. Thus we may expect school districts who will benefit from increased property values to provide more incentives for ratings improvement. Thus, future versions of this paper will explore the interaction between school finance equalization and the accountability system.

References

Carnoy, Martin and Susanna Loeb. "Does External Accountability Affect Student Outcomes? A Cross-State Analysis." *Education Evaluation and Policy Analysis*, 24, 2003, pp. 305 - 331.

Cullen, Julie Berry and Randall Reback. "Tinkering Toward Accolades: School Gaming under a Performance Accountability System." NBER Working Paper #12286, June 2006.

Chakrabarti, Rajashri. "Impact of Voucher Design on Public School Performance: Evidence from Florida and Milwaukee Voucher Programs." Federal Reserve Bank of New York Staff Report no. 315, January 2008.

Figlio, David and Cecelia Elana Rouse. "Do Accountability and Voucher Threats Improve Low-Performing Schools?" *Journal of Public Economics*, 90, 2006, pp. 239 - 255.

---. "Testing, Crime, and Punishment." *Journal of Public Economics*, 90, 2006, pp. 837 - 851.

--- and Joshua Winicki. "Food for Thought: The Effects of School Accountability Plans on School Nutrition." *Journal of Public Economics*, 89, 2005, pp. 381 - 394.

--- and Maurice E. Lucas. "What's in a Grade? School Report Cards and the Housing Market." *American Economic Review*, 94, June 2004, pp. 591 - 604.

--- and Lawrence S. Getzler. "Accountability, Ability, and Disability: Gaming the System." NBER Working Paper #9307, October, 2002.

Hahn, Jinyong, Petra Todd, and Wilbert Van Der Klaauw, "Identification and Estimation of Treatment Effects with a Regression-discontinuity Design," *Econometrica*, 69, January 2001, pp 201-09.

Hanushek, Eric A. and Margaret E. Raymond. "Does School Accountability Lead to Improved Student Performance?" *Journal of Policy Analysis and Management*, 24, 2005, pp. 297 - 327.

--- and ---. "The Effect of School Accountability Systems on the Level and Distribution of Student Achievement." *Journal of the European Economic Association*, 2, April-May 2004, pp. 406 - 415.

Jacob, Brian A. "Accountability, Incentives, and Behavior: The Impact of High-Stakes Testing in the Chicago Public Schools." *Journal of Public Economics*, 89, 2005, pp. 761 - 796.

---. "Getting Inside Accountability: Lessons from Chicago." *Brookings-Wharton Papers on Urban Affairs*, 2003, pp. 42 - 70.

Kane, Thomas J., Douglas O. Staiger and Gavin Samms. "School Accountability Ratings and Housing Values." Brookings-Wharton Papers on Urban Affairs, 2003, pp. 83 - 127.
Neal, Derek and Diane Whitmore Schanzenbach. "Left Behind by Design: Proficiency Counts and Test-Based Accountability." NBER Working Paper #13293, August 2007.

Rebeck, Randall. "Teaching to the Rating: School Accountability and the Distribution of Student Achievement." *Journal of Public Economics*, 92, 2008, pp. 1394 - 1415.

Rouse, Cecelia Elena, Jane Hannaway, Dan Goldhaber and David Figlio. "Feeling the Florida Heat? How Low-Performing Schools Respond to Voucher and Accountability Pressure." NBER Working Paper #13681, December 2007.

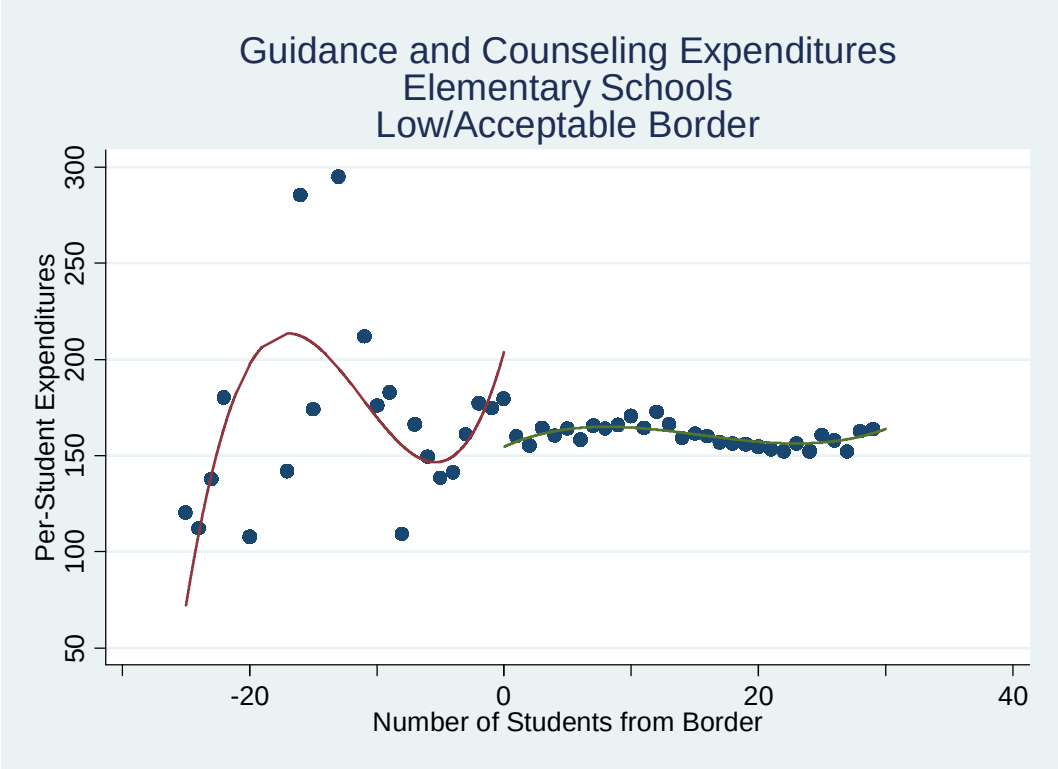

Figure 1: Elementary School Re-Allocation via Guidance and Counseling Expenditures - Low/Acceptable

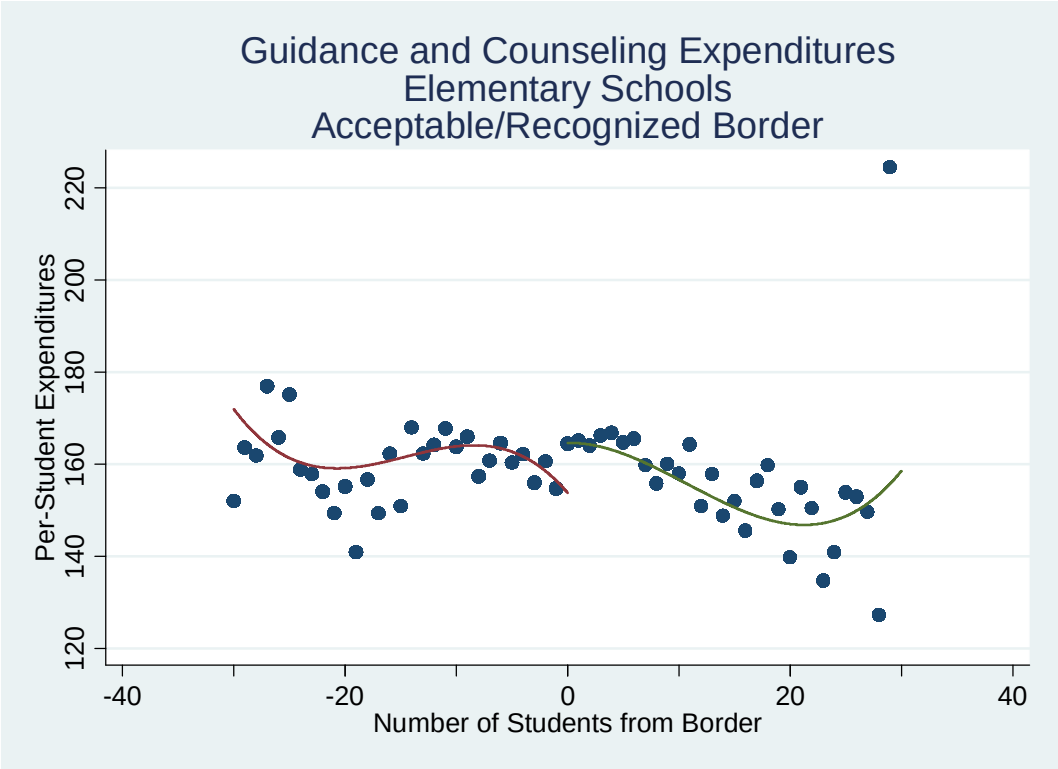

Figure 2: Elementary School Re-Allocation via Guidance and Counseling Expenditures - Acceptable/Recognized

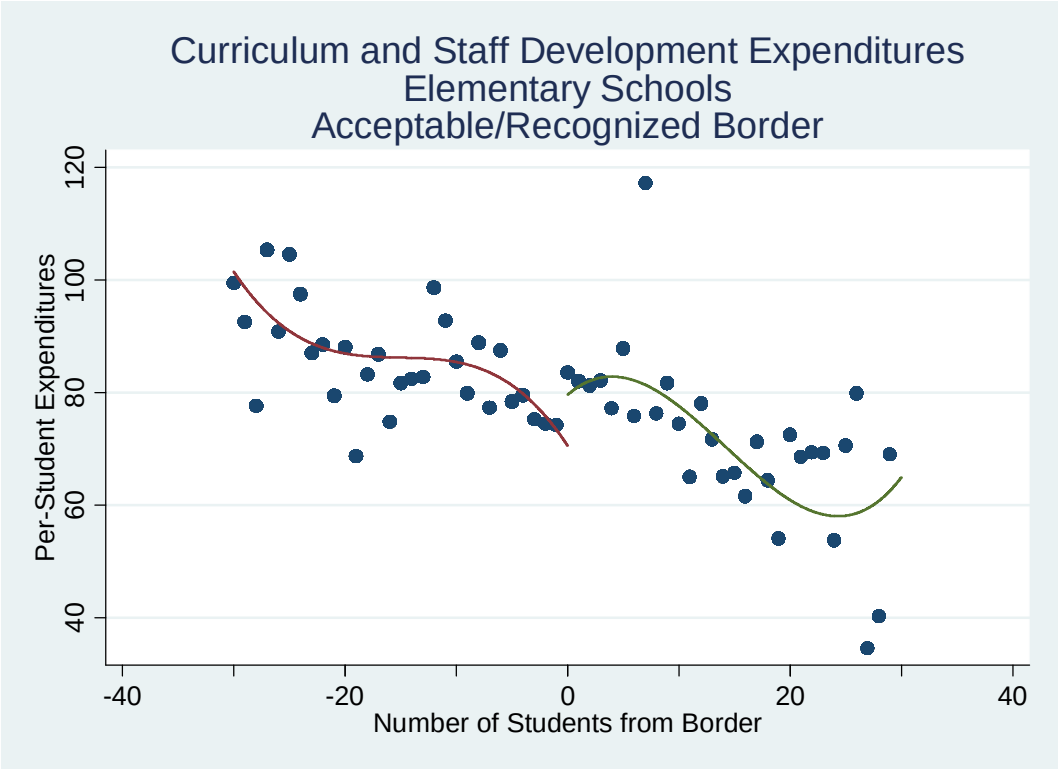

Figure 3: Elementary School Re-Allocation via Curriculum and Staff Development Expenditures - Acceptable/Recognized

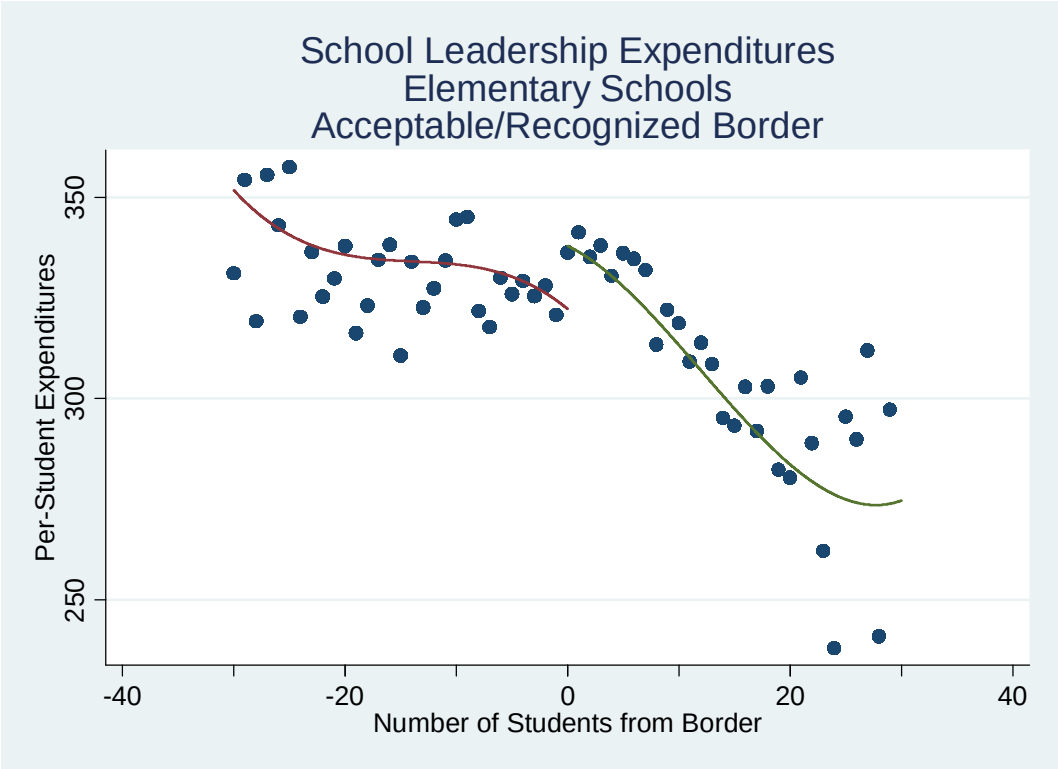

Figure 4: Elementary School Re-Allocation via School Leadership Expenditures - Acceptable/Recognized

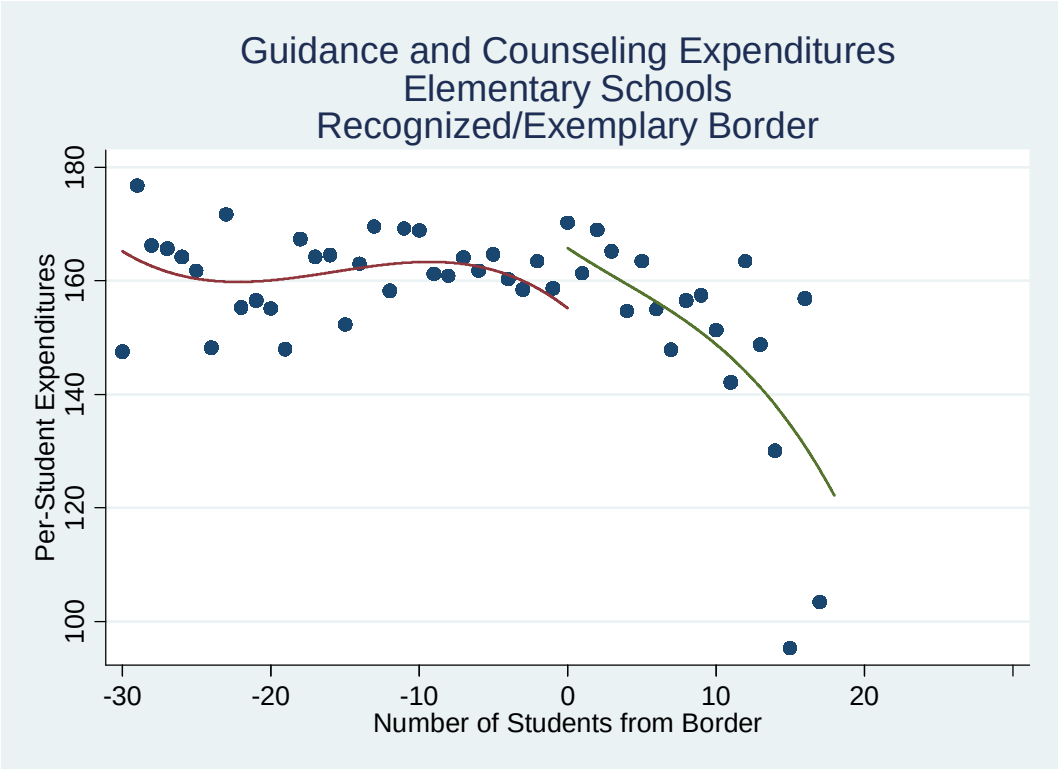

Figure 5: Elementary School Re-Allocation via Guidance & Counseling Expenditures - Recognized/Exemplary

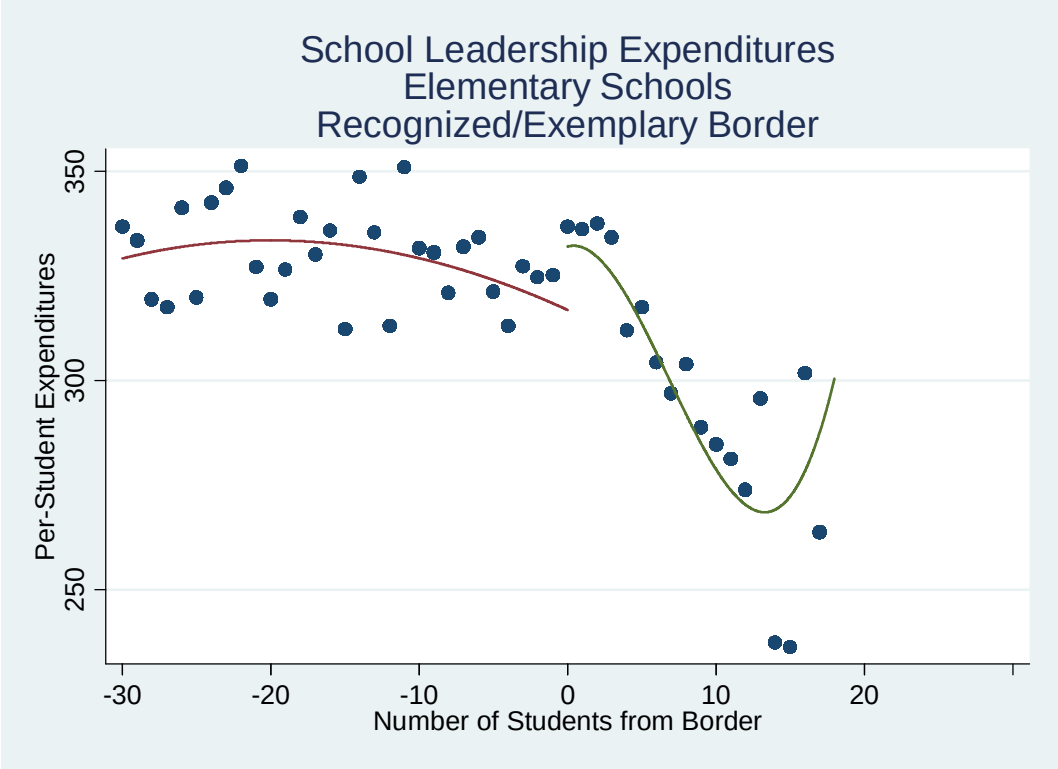

Figure 6: Elementary School Re-Allocation via School Leadership Expenditures - Recognized/Exemplary

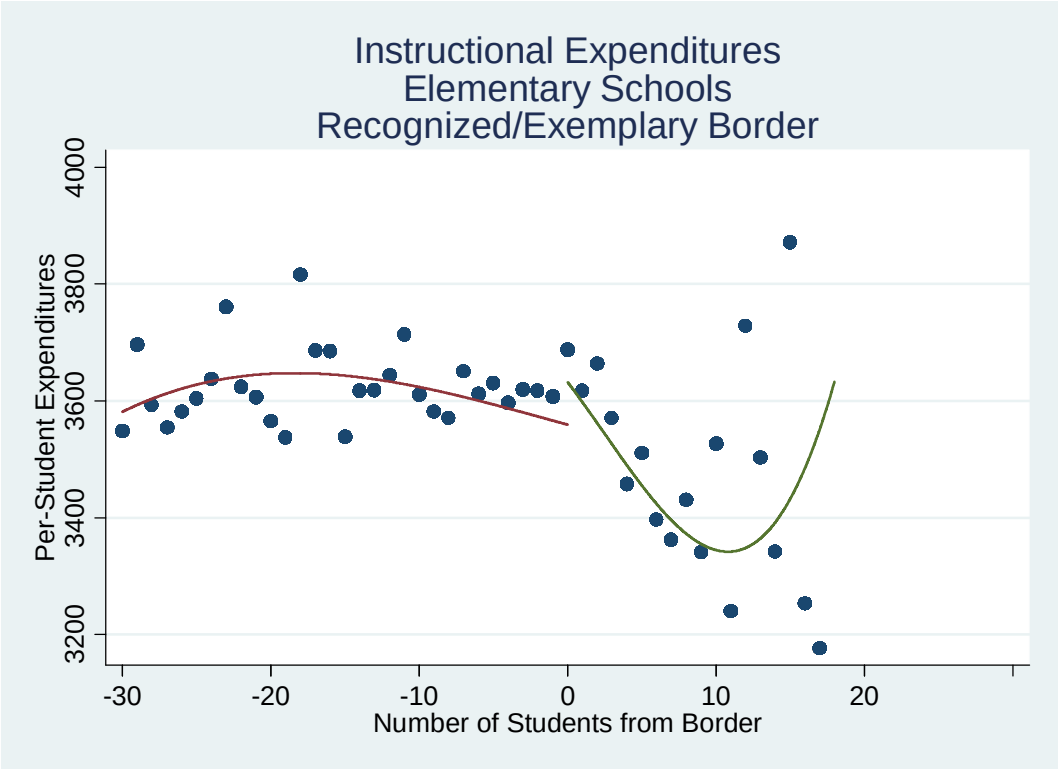

Figure 7: Elementary School Re-Allocation via Instructional Expenditures - Recognized/Exemplary

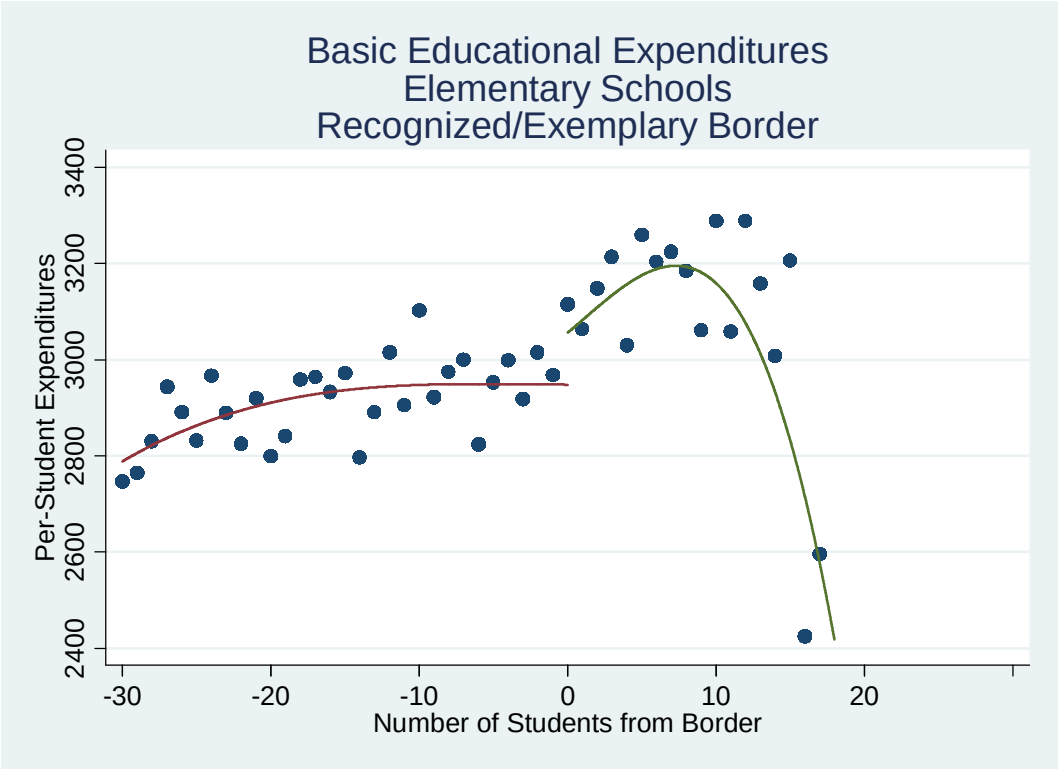

Figure 8: Elementary School Re-Allocation via Basic Educational Expenditures - Recognized/Exemplary

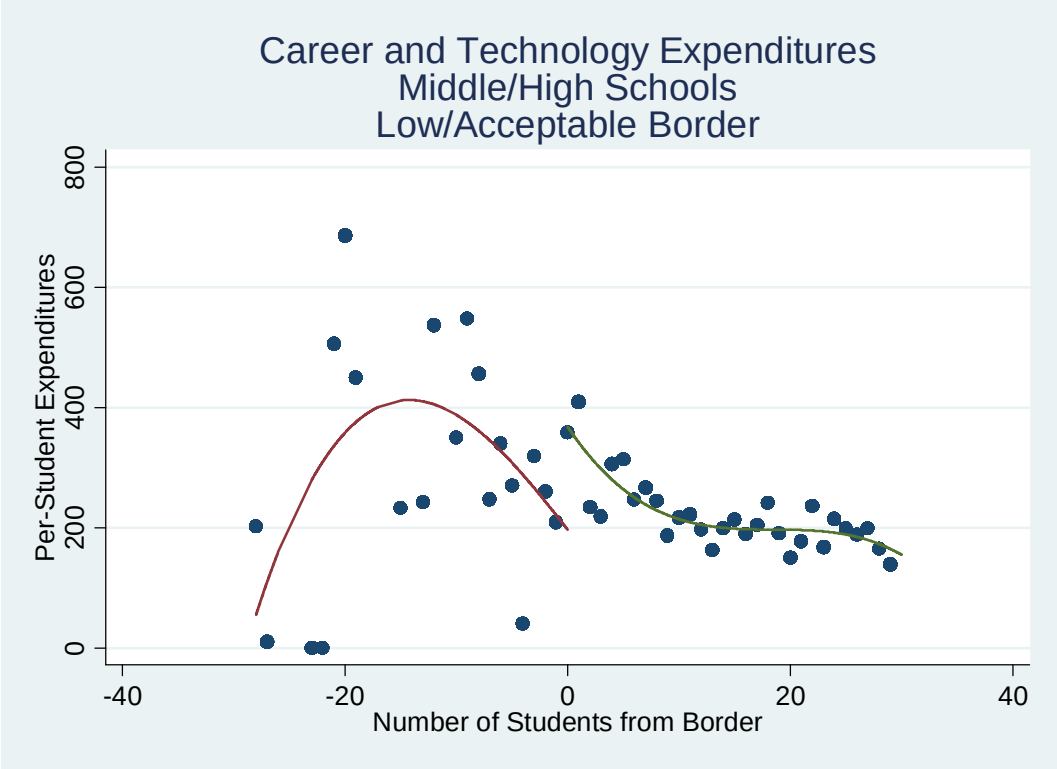

Figure 9: Middle and High School Re-Allocations via Technology Expenditures

Table 1: Texas TAAS Accountability Rules

1998-2002

From Low Performing to Acceptable

	Math, Reading	Writing	Social Studies	Drop-Outs	Attendance
Grades	3 - 8, 10	4, 8, 10	8	7 - 12	All
Groups	White, Black, Hisp, Econ Dis, All	White, Black, Hisp, Econ Dis, All	All, only	White, Black, Hisp, Econ Dis, All	White, Black, Hisp, Econ Dis, All
1998-99	40%	40%	n/a	6%	94%
1999-00	45%	45%	n/a	6%	94%
2000-01	50%	50%	n/a	5.5%	94%
2001-02	50%	50%	n/a	5%	-
2002-03	55%	55%	50%	5%	-

Acceptable/Recognized

	Math, Reading	Writing	Social Studies	Drop-Outs	Attendance
Grades	3 - 8, 10	4, 8, 10	8	7 - 12	All
Groups	White, Black, Hisp, Econ Dis, All	White, Black, Hisp, Econ Dis, All	All, only	White, Black, Hisp, Econ Dis, All	White, Black, Hisp, Econ Dis, All
1998-99	80%	80%	n/a	3.5%	94%
1999-00	80%	80%; 3-8, 10	n/a	3.5%	94%
2000-01	80%	80%; 3-8, 10	n/a	3.0%	94%
2001-02	80%	80%; 3-8, 10	n/a	2.5%	-
2002-03	80%	80%; 3-8, 10	80%	2.5%	-

Recognized/Exemplary

	Math, Reading	Writing	Social Studies	Drop-Outs	Attendance
Grades	3 - 8, 10	4, 8, 10	8	7 - 12	All
Groups	White, Black, Hisp, Econ Dis, All	White, Black, Hisp, Econ Dis, All	All, only	White, Black, Hisp, Econ Dis, All	White, Black, Hisp, Econ Dis, All
1998-99	90%	90%	n/a	1%	94%
1999-00	90%	90%; 3-8, 10	n/a	1%	94%
2000-01	90%	90%; 3-8, 10	n/a	1%	94%
2001-02	90%	90%; 3-8, 10	n/a	1%	-
2002-03	90%	90%; 3-8, 10	90%	1%	-

Notes: To count, all of the subject/student group combinations must be at least either: 30 students, or 10% of the student body, or 200 students. If not, then special analysis can be used, which may include prior year data.

Table 2: Sample Means by School Ranking

		School Characteristics by Ranking 1997-98 through 2001-02			
		Low	Acceptable	Recognized	Exemplary
Expenditures by Function (per-pupil)	Total Expenditures	5,886	5,308	5,432	5,782
		(3494)	(2301)	(2073)	(2958)
	Instruction	3,809	3,500	3,626	3,742
		(2758)	(1039)	(810)	(1032)
	Instructional Resources	120	125	134	145
		(76)	(68)	(74)	(80)
	Curriculum & Staff Development	103	69	64	63
		(153)	(87)	(101)	(269)
	Instructional Leadership	81	66	59	54
		(58)	(55)	(54)	(51)
	School Leadership	380	344	346	358
		(228)	(180)	(126)	(147)
	Guidance & Counseling	197	171	163	163
		(134)	(106)	(82)	(90)
	Social Work Services	19	12	10	7
		(51)	(25)	(18)	(19)
	Co-/Extra Curricular Activities	94	104	115	145
		(160)	(200)	(215)	(284)
Expenditures by Intended Program (per-pupil)	Basic Educational Services	2,686	2,797	2,970	3,218
		(846)	(868)	(784)	(888)
	Career & Tech	161	143	132	173
		(281)	(519)	(288)	(345)
Staff - Student Ratios	Teachers	0.068	0.070	0.071	0.073
		(0.022)	(0.018)	(0.015)	(0.017)
	Educational Aides	0.013	0.015	0.016	0.015
		(0.011)	(0.011)	(0.010)	(0.011)
	Administrators	0.004	0.004	0.004	0.004
		(0.003)	(0.002)	(0.002)	(0.002)
School Characteristics	% Asian/Pacific Islander	1	2	2	3
		(2)	(3)	(4)	(5)
	% Black	30	18	11	7
		(27)	(22)	(16)	(11)
	% Hispanic	48	43	38	24
		(30)	(32)	(32)	(28)
	% White	21	37	49	66
		(23)	(30)	(31)	(28)
	% Economically Disadvantaged	71	60	51	33
		(23)	(25)	(26)	(27)
	% LEP (Limited English Proficiency)	23	16	12	7

Table 2: Sample Means by School Ranking

	School Characteristics by Ranking 1997-98 through 2001-02			
	Low	Acceptable	Recognized	Exemplary
	(23)	(19)	(17)	(14)
% Special Education	18	15	13	15
	(26)	(24)	(24)	(28)
% Vocational	12	13	13	13
	(8)	(5)	(5)	(5)
Enrollment	779	705	586	537
	(515)	(510)	(415)	(389)
Observations (Expenditures & School Characteristics)	303	13,119	9,563	6,479
Observations (Staff - Student Ratios)	300	13,111	9,559	6,478

Table 2A: Percent of Schools with Same Rating in Two Different Years

I. Low _Performing - 231 School-Years

		School is Low in				
		1997	1998	1999	2000	2001
% of Schools Low in	1997	-	0.0	3.9	0.0	1.6
	1998	0.0	-	6.4	3.9	4.8
	1999	9.4	15.6	-	26.9	17.5
	2000	0.0	3.1	8.9	-	4.8
	2001	3.1	9.4	14.1	11.5	-

II. Acceptable - 9,252 School-Years

		School is Acceptable in				
		1997	1998	1999	2000	2001
% of Schools Acceptable in	1997	-	78.2	74.5	72.7	73.6
	1998	75.3	-	77.1	74.8	73.2
	1999	68.0	73.1	-	76.7	74.3
	2000	55.9	59.7	64.5	-	72.2
	2001	43.0	44.4	47.6	54.9	-

III. Recognized - 6,207 School-Years

		School is Recognized in				
		1997	1998	1999	2000	2001
% of Schools Recognized in	1997	-	46.0	39.8	32.6	28.3
	1998	50.5	-	47.7	37.9	31.5
	1999	49.6	54.1	-	46.2	40.5
	2000	50.7	53.7	57.7	-	51.7
	2001	43.9	44.4	50.4	51.4	-

IV. Exemplary - 4,143 School-Years

		School is Exemplary in				
		1997	1998	1999	2000	2001
% of Schools Exemplary in	1997	-	67.3	60.2	49.6	42.2
	1998	68.3	-	68.3	53.9	44.7
	1999	63.7	71.2	-	60	50.7
	2000	65.7	70.3	75.1	-	60.3
	2001	67.4	70.5	76.6	72.8	-

Table 3: OLS Impact Estimates of Accountability on Resources

I. Elementary Schools

	Acceptable	Recognized	Exemplary	Observations
Total Expenditures	-183.1	-219.7*	-377.9***	13,748
	(122.3)	(126.0)	(134.1)	
Instruction	-192.2**	-215.9***	-330.3***	13,748
	(76.4)	(78.2)	(80.1)	
Instructional Resources	0.7	3.3	3.7	13,748
	(5.0)	(5.2)	(5.3)	
Curriculum & Staff Development	-46.2***	-48.8***	-51.2***	13,748
	(15.5)	(16.5)	(19.3)	
Instructional Leadership	1.6	0.7	-3.0	13,748
	(2.0)	(2.1)	(2.2)	
School Leadership	-17.8	-17.9	-20.0	13,748
	(12.1)	(12.6)	(12.8)	
Guidance & Counseling	-11.6*	-10.8*	-14.8**	13,748
	(6.2)	(6.3)	(6.6)	
Social Work Services	-0.3	-2.5**	-4.1***	13,748
	(1.1)	(1.1)	(1.2)	
Co-/Extra Curricular Activities	1.1	1.1	2.0**	13,748
	(0.9)	(0.9)	(1.0)	
Basic Educational Services	-109.0	-11.2	103.5	13,748
	(77.4)	(79.0)	(80.5)	
Staff - Student Ratios				
Teachers	0.00052	0.00016	-0.00062	13,745
	(0.00103)	(0.00106)	(0.00109)	
Educational Aides	-0.00052	-0.00078	-0.00190***	13,745
	(0.00054)	(0.00056)	(0.00058)	
Administrators	-0.00024	-0.00031	-0.00037	13,745
	(0.00022)	(0.00023)	(0.00023)	

Notes: Each category is regressed separately on the rating dummy variables, fixed effects for school district, and for years. We report the coefficient estimates on the rating dummy variables.

* indicates significance at the 10% level, ** indicates significance at the 5% level, and *** indicates significance at the 1% level.
Standard errors are in parentheses, clustered by school.

Table 3: OLS Impact Estimates of Accountability on Resources (cont)

II. Pooled Middle and High Schools

	Acceptable	Recognized	Exemplary	Observations
Total Expenditures	-735.0***	-1075.9***	-992.5***	5,639
	(284.2)	(293.6)	(321.8)	
Instruction	-258.9	-354.8*	-434.0**	5,639
	(206.4)	(209.4)	(213.6)	
Instructional Resources	-3.6	-6.0	-7.5	5,639
	(4.1)	(4.5)	(5.1)	
Curriculum & Staff Development	-16.7**	-22.4***	-24.2***	5,639
	(7.7)	(7.7)	(7.7)	
Instructional Leadership	-5.8	-9.9**	-14.7***	5,639
	(4.4)	(4.6)	(4.9)	
School Leadership	-18.9	-27.0	-25.8	5,639
	(16.0)	(17.1)	(18.3)	
Guidance & Counseling	-8.9	-22.2**	-32.6***	5,639
	(8.2)	(8.8)	(10.1)	
Social Work Services	-3.3	-5.8**	-7.0***	5,639
	(2.2)	(2.3)	(2.3)	
Co-/Extra Curricular Activities	-40.2***	-55.1***	-51.7***	5,639
	(11.1)	(11.9)	(13.2)	
Basic Educational Services	69.8	61.2	78.7	5,639
	(71.6)	(79.2)	(90.2)	
Career & Tech	-44.6	-48.8	-5.0	5,639
Staff-Student Ratios	(27.8)	(33.7)	(43.7)	
Teachers	0.00275*	0.00151	0.00113	5,636
	(0.00154)	(0.00159)	(0.00169)	
Educational Aides	0.00022	-0.00005	-0.00117**	5,636
	(0.00049)	(0.00052)	(0.00060)	
Administrators	0.00002	-0.00015	-0.00024	5,636
	(0.00021)	(0.00022)	(0.00022)	

Notes: Each category is regressed separately on the rating dummy variables, fixed effects for school district, and for years. We report the coefficient estimates on the rating dummy variables.
* indicates significane at the 10% level, ** indicates significance at the 5% level, and *** indicates significance at the 1% level.
Standard errors are in parentheses, clustered by school.

Table 4: School District Re-Allocations Among Elementary Schools

	Means (Std Dev)	Low-Acceptable		Acceptable-Recognized		Recognized-Exemplary	
		(1)	(2)	(3)	(4)	(5)	(6)
Total Expenditures	5362	-452.7	103.8	218.4*	116.1	57.2	94.8
	(1848)	(396.3)	(383.2)	(115.3)	(94.2)	(160.3)	(131.1)
Instruction	3587	-498.1**	-284.1	53.4	43.2	57.6	67.3*
	(660)	(240.3)	(211.5)	(43.7)	(36.2)	(50.0)	(40.2)
Basic Educational Services	2849	143.6	72.9	-52.0	-55.4	84.9*	103.0**
	(733)	(202.0)	(180.8)	(48.4)	(40.4)	(50.4)	(41.5)
Instructional Resources	134	-26.6	-15.1	4.5	5.2	7.1	4.5
	(60)	(20.8)	(17.4)	(4.4)	(3.7)	(4.5)	(3.7)
Curriculum & Staff Development	86	24.1	64.2	10.2*	9.0*	38.4	29.0
	(204)	(46.8)	(47.7)	(5.6)	(4.8)	(30.3)	(23.3)
Instructional Leadership	77	-3.6	0.8	1.3	2.5	3.1	5.8**
	(50)	(8.5)	(7.7)	(4.1)	(3.4)	(3.8)	(2.9)
School Leadership	329	-36.6	-20.1	18.1***	15.6***	11.8	14.6**
	(103)	(35.5)	(32.3)	(6.7)	(5.7)	(7.8)	(6.3)
Guidance & Counseling	160	-60.1**	-49.1**	10.1*	10.8***	8.6*	10.3**
	(68)	(25.0)	(21.9)	(5.3)	(4.1)	(5.1)	(4.3)
Social Work Services	13	1.0	2.5	0.3	0.7	-0.5	0.8
	(20)	(4.6)	(4.2)	(1.4)	(1.2)	(1.3)	(1.2)
Co-/Extra Curricular Activities	10	5.9**	4.5*	-2.5	-0.8	-0.5	-0.5
	(25)	(2.6)	(2.7)	(1.6)	(1.3)	(2.1)	(1.8)
Staff-Student Ratios							
Teachers	15.5	-0.0087**	-0.0051	0.0007	0.0008	0.0000	0.0006
	(6.4)	(0.0038)	(0.0033)	(0.0007)	(0.0006)	(0.0008)	(0.0006)
Educational Aides	89.2	-0.0058*	-0.0036	-0.0002	-0.0005	0.0019***	0.0012**
	(103.9)	(0.0030)	(0.0025)	(0.0006)	(0.0005)	(0.0007)	(0.0005)
Administrators	356.1	-0.0005	-0.0004	0.0000	0.0001	0.0001	0.0002**
	(148.7)	(0.0005)	(0.0005)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Range		-20 to 20	-30 to 30	-20 to 20	-30 to 30	-20 to 20	-30 to 30
Obs Above Border (Functional Exp)		9329	11914	7764	7912	3289	3289
Obs Below Border (Functional Exp)		122	127	2808	3312	3477	4254

Notes: These are the intercept terms from a regression with the marginal defined as by Range, separately for each of the three borders.

* indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Regressions use all school districts with more than one elementary school, where the school size is at least 200 students.

Standard errors are in parentheses, clustered by school.

The sample excludes alternative schools, charters, unrated schools, and schools given ratings on appeal.

Table 5: School District Re-Allocations Among Pooled Middle and High Schools

	Means (Std Dev)	Low-Acceptable		Acceptable-Recognized		Recognized-Exemplary	
		(1)	(2)	(3)	(4)	(5)	(6)
Total Expenditures	5645	-32.7	778.0	82.9	-26.7	82.9	-159.4
	(2239)	(760.1)	(692.7)	(249.5)	(193.7)	(436.5)	(312.8)
Instruction	3520	-5.4	-53.8	94.1	100.8	31.4	15.6
	(915)	(202.9)	(178.2)	(80.7)	(65.4)	(132.6)	(109.0)
Basic Educational Services	3102	-109.8	-124.9	38.2	27.8	67.5	2.7
	(850)	(166.2)	(148.5)	(68.4)	(55.4)	(161.0)	(128.0)
Instructional Resources	118	-11.2	-10.8	6.5	4.5	1.4	-2.1
	(53)	(15.6)	(12.0)	(6.0)	(5.1)	(8.6)	(7.2)
Curriculum & Staff Development	74	-37.9	-42.9*	-1.3	-1.2	3.6	3.7
	(80)	(29.7)	(23.6)	(6.6)	(5.1)	(9.7)	(8.2)
Instructional Leadership	76	5.2	3.9	4.2	0.0	2.9	-0.2
	(41)	(15.4)	(11.4)	(4.7)	(3.8)	(5.7)	(4.4)
School Leadership	381	-21.4	-15.7	18.1	28.1	-0.9	6.1
	(180)	(53.2)	(45.7)	(20.4)	(17.6)	(23.1)	(18.8)
Guidance & Counseling	231	-23.2	-13.6	-0.1	1.9	-5.1	0.1
	(91)	(28.1)	(22.3)	(10.9)	(8.6)	(16.1)	(12.1)
Social Work Services	15	7.9*	-1.2	1.0	0.4	0.7	1.3
	(23)	(4.8)	(5.3)	(2.7)	(1.9)	(3.0)	(2.3)
Co-/Extra Curricular Activities	128	35.5	3.7	6.8	9.0	18.3	-4.1
	(123)	(39.8)	(37.5)	(14.0)	(11.3)	(20.2)	(14.7)
Career & Tech	209	178.9*	172.3**	-11.1	21.9	48.8	61.5
	(721)	(105.4)	(83.4)	(34.5)	(29.3)	(87.6)	(62.6)
Staff - Student Ratios							
Teachers	15.1	0.0100*	0.0095**	0.0001	0.0005	0.0009	-0.0008
	(6.7)	(0.0059)	(0.0045)	(0.0014)	(0.0011)	(0.0019)	(0.0015)
Educational Aides	159.9	0.0002	0.0014	-0.0002	0.0001	-0.0001	-0.0008
	(235.1)	(0.0018)	(0.0015)	(0.0008)	(0.0007)	(0.0010)	(0.0008)
Administrators	317.5	0.0004	0.0008	0.0000	0.0000	-0.0002	-0.0002
	(126.0)	(0.0009)	(0.0007)	(0.0002)	(0.0001)	(0.0002)	(0.0002)
Range		-20 to 20	-30 to 30	-20 to 20	-30 to 30	-20 to 20	-30 to 30
Obs Above Border (Functional Exp)		2549	3683	1744	1845	587	659
Obs Below Border (Functional Exp)		72	77	1170	1427	751	1005

Notes: These are the intercept terms from a regression with the marginal defined as by Range, separately for each of the three borders. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.
Regressions use all school districts with more than one middle or high school, where the school size is at least 200 students.
Standard errors are in parentheses, clustered by school.
The sample excludes alternative schools, charters, unrated schools, and schools given ratings on appeal.

Table 6: School District Re-Allocations Among Elementary Schools

Using the Five Largest School Districts in Texas

	Low-Acceptable		Acceptable-Recognized		Recognized-Exemplary	
	(1)	(2)	(3)	(4)	(5)	(6)
Total Expenditures	-500.0	415.2	560.3*	594.3**	1478.8	1311.3
	(486.3)	(507.1)	(328.7)	(293.1)	(1116.1)	(959.7)
Instruction	-502.5*	-222.9	164.3	218.7*	165.0	204.5
	(280.8)	(258.7)	(138.1)	(123.5)	(243.2)	(182.6)
Instructional Resources	-40.3*	-20.9	11.4	20.8**	5.5	6.1
	(24.3)	(22.0)	(10.8)	(9.3)	(17.0)	(13.6)
Curriculum & Staff Development	23.3	121.4*	13.5	6.3	375.6	300.4
	(71.0)	(66.3)	(27.2)	(27.0)	(361.0)	(281.8)
Instructional Leadership	-12.2	-10.5	16.4**	15.0**	-3.2	0.1
	(8.7)	(7.7)	(7.3)	(6.2)	(10.7)	(8.4)
School Leadership	-38.9	-7.1	26.6	34.2	24.9	29.0
	(45.8)	(41.9)	(23.6)	(21.4)	(45.2)	(34.8)
Guidance & Counseling	-56.5**	-39.3	35.5**	30.5**	11.2	9.2
	(27.0)	(25.1)	(16.6)	(13.8)	(21.8)	(16.6)
Social Work Services	5.7	5.7*	0.0	2.0	-4.0	-2.3
	(3.5)	(3.1)	(3.1)	(3.1)	(4.6)	(3.5)
Co-/Extra Curricular Activities	3.6	3.8	-1.6	-1.7	0.0	0.0
	(2.8)	(3.0)	(1.2)	(1.2)	(1.1)	(0.5)
Basic Educational Services	103.0	214.6	152.0	71.0	34.2	1.9
	(246.7)	(229.7)	(160.5)	(143.1)	(241.6)	(180.7)
Staff-Student Ratios						
Teachers	-0.0054	-0.0009	0.0025	0.0036**	-0.0058*	-0.0020
	(0.0034)	(0.0032)	(0.0019)	(0.0017)	(0.0035)	(0.0027)
Educational Aides	-0.0053**	-0.0045**	0.0012	0.0011	0.0024	0.0024*
	(0.0026)	(0.0023)	(0.0012)	(0.0010)	(0.0017)	(0.0013)
Administrators	-0.0005	-0.0003	0.0001	0.0003	-0.0003	0.0000
	(0.0007)	(0.0007)	(0.0003)	(0.0002)	(0.0004)	(0.0003)
Range	-20 to 20	-30 to 30	-20 to 20	-30 to 30	-20 to 20	-30 to 30
Obs Above Border (Functional Exp)	1985	2379	884	889	258	258
Obs Below Border (Functional Exp)	76	80	452	582	348	464

Notes: These are the intercept terms from a regression with the marginal defined as by Range, separately for each of the three borders. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.
Regressions use all school districts with more than one elementary school, where the school size is at least 200 students, in the five largest school districts in Texas, which are Houston, Dallas, Fort Worth, Austin, and San Antonio.
Standard errors are in parentheses, clustered by school.
The sample excludes alternative schools, charters, unrated schools, and schools given ratings on appeal.

Table 7: School District Re-Allocations Among Pooled Middle and High Schools

Using the Five Largest School Districts in Texas

		Low-Acceptable		Acceptable-Recognized		Recognized-Exemplary	
	Means (Std Dev)	(1)	(2)	(3)	(4)	(5)	(6)
Total Expenditures	6078	-82.5	1252.5	-477.4	-348.0	1103.9	1.2
	(2596)	(1309.1)	(1116.5)	(1105.1)	(601.8)	(2338.4)	(2017.4)
Instruction	3620	260.2	4.8	-398.8	-93.7	563.4	22.3
	(1284)	(289.9)	(218.8)	(464.9)	(312.2)	(1349.3)	(1222.1)
Instructional Resources	106	-0.6	-7.4	9.3	2.1	13.8	-18.3
	(53)	(19.0)	(13.9)	(14.9)	(14.6)	(53.3)	(40.3)
Curriculum & Staff Development	93	-44.4	-49.8	-35.4	-15.8	77.1	33.3
	(112)	(40.7)	(30.5)	(39.9)	(25.0)	(53.9)	(55.6)
Instructional Leadership	79	7.6	-5.7	-23.2*	-26.6***	29.3	25.5
	(46)	(18.8)	(15.4)	(12.3)	(10.1)	(20.4)	(18.7)
School Leadership	418	26.6	4.6	-27.2	39.0	84.3	-13.4
	(227)	(92.8)	(75.3)	(104.0)	(74.0)	(195.9)	(179.4)
Guidance & Counseling	251	-14.4	-17.4	-11.0	-8.7	74.1	-35.4
	(107)	(40.4)	(27.8)	(50.0)	(37.3)	(113.8)	(91.5)
Social Work Services	17	13.4	-4.5	-23.0	-10.7	14.8*	11.4
	(28)	(9.0)	(7.9)	(14.4)	(12.1)	(8.7)	(8.2)
Co-/Extra Curricular Activities	182	-4.9	8.7	-56.3	-21.0	-24.7	15.1
	(168)	(46.8)	(39.2)	(41.7)	(33.0)	(38.7)	(35.2)
Basic Educational Services	3005	-104.2	-94.7	-105.1	8.5	407.0	-205.7
	(967)	(222.6)	(171.6)	(311.6)	(248.4)	(1970.2)	(1578.4)
Career & Tech	456	20.3	142.6	-115.5	8.5	-295.3	79.0
	(779)	(156.6)	(115.5)	(173.8)	(158.8)	(834.5)	(557.8)
Staff - Student Ratios							
Teachers	0.066	0.0058	0.0016	-0.0037	-0.0022	0.0105	0.0068
	(0.029)	(0.0038)	(0.0032)	(0.0046)	(0.0034)	(0.0099)	(0.0087)
Educational Aides	0.007	0.0007	-0.0008	-0.0019	-0.0024	0.0060**	0.0025
	(0.005)	(0.0023)	(0.0017)	(0.0023)	(0.0015)	(0.0028)	(0.0029)
Administrators	0.004	-0.0002	-0.0006	-0.0002	-0.0001	0.0003	-0.0007
	(0.002)	(0.0009)	(0.0007)	(0.0006)	(0.0004)	(0.0008)	(0.0008)
Range		-20 to 20	-30 to 30	-20 to 20	-30 to 30	-20 to 20	-30 to 30
Obs Above Border (Functional Exp)		502	651	148	150	43	43
Obs Below Border (Functional Exp)		39	42	94	121	44	58

Notes: These are the intercept terms from a regression with the marginal defined as by Range, separately for each of the three borders. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.
Regressions use all school districts with more than one middle or high school, where the school size is at least 200 students, in the five largest school districts in Texas, which are Houston, Dallas, Fort Worth, Austin, and San Antonio.
Standard errors are in parentheses, clustered by school.
The sample excludes alternative schools, charters, unrated schools, and schools given ratings on appeal.