Program on Education Policy and Governance Working Papers Series

Does Information Help Families Choose Schools? Evidence from a Regression Discontinuity Design

Michael Henderson
PhD Candidate
Government and Social Policy
Harvard University

PEPG 10-17

Program on Education Policy and Governance
Harvard Kennedy School
79 JFK Street, Taubman 304
Cambridge, MA 02138
Tel: 617-495-7976 Fax: 617-496-4428
www.hks.harvard.edu/pepg/

This research has been supported by the National Science Foundation's IGERT program,\emph{Multi-disciplinary Program in Inequality \& Social Policy} at Harvard University (Grant No. 0333403) as well as by the Program on Education Policy and Governance at Harvard University. Thanks to Matthew Chingos, Paul Peterson, Eric Hanushek, Bruce Western, Kathy Edin, William Julius Wilson, and participants in the Proseminar on Inequality and Social Policy at Harvard University for helpful comments. An earlier draft of this paper was presented at the 2009 conference of the National Center for School Choice at Vanderbilt University and appeared as PEPG 09-06 in the working papers series.

Does Information Help Families Choose Schools? Evidence from a Regression Discontinuity Design*

August 4, 2010

Michael Henderson

PhD Candidate

Government and Social Policy

Harvard University

1737 Cambridge St

Cambridge, MA 02138

Abstract

Does public information about school quality lead parents to sort their children out of schools with relatively poor performance? Use of this exit option in response to information about school quality has the potential to indirectly foster school responsiveness to quality concerns. To determine whether this information affects student exit, I use a regression discontinuity design to examine the effect of school grades on exit. Results indicate that parents do not seem to respond to information about school quality generally and, thus, cast doubt on the effectiveness of indirect accountability to promote educational improvement. However, there is limited evidence that particularly poor school performance accompanied by institutional mechanisms for school choice promote student sorting away from low-quality schools.

*This research has been supported by the National Science Foundation's IGERT program, \emph{Multidisciplinary Program in Inequality \& Social Policy} at Harvard University (Grant No. 0333403) as well as by the Program on Education Policy and Governance at Harvard University. Thanks to Matthew Chingos, Paul Peterson, Eric Hanushek, Bruce Western, Kathy Edin, William Julius Wilson, and participants in the Proseminar on Inequality and Social Policy at Harvard University for helpful comments. A draft of this paper was presented at the 2009 conference of the National Center for School Choice at Vanderbilt University.

I. Accountability, Information, and Choice

In recent decades accountability has been the most politically successful school reform proposal in the United States. During the 1990s 31 states launched accountability programs, and with passage of the federal No Child Left Behind Act in 2002 receipt of federal education dollars became contingent upon adoption of a U.S. Department of Education approved system of annual testing and performance-based accountability. Accountability found consensus even as other reforms such school choice have remained bogged in the mud of political battles. While supporters of choice have suggested that the accountability approach and the choice approach can be mutually reinforcing, opponents of choice have frequently argued for accountability as an alternative to more radical choice-based reforms. So while accountability programs across the U.S. include small scale public school choice components, only rarely have these state programs incorporated private school choice. Florida in the early 2000s, however, is one such case.

Yet, the effectiveness of accountability seems to depend in part on the logic of choice in response to information. Across the fifty states, accountability programs share an essential element: public-release of school performance results. In the most direct form of accountability, the schools face specific sanctions based on performance results as prescribed in accountability laws. But these sanctions are typically reserved only for schools that consistently fail to meet minimum performance standards. With only limited implementation of these direct incentives, the prospects for accountability to improve school quality depends on indirect mechanisms. At all performance levels schools face the release of quality reports to the public at large and parents specifically. The public has recourse to the ballot box for rewarding or punishing school governing officials, and parents have recourse to their feet - i.e., to use this information in decisions about where to enroll their children.

Indirect accountability operates outside the specific system of sanctions and rewards built into a state's program. Along with direct accountability, this notion shares the idea that schools will respond to an incentive structure built around feedback from their performance. The formal structure provides information while the "accountability" is exercised voluntarily by stakeholders. When a person becomes dissatisfied with a particular institution (her grocery store or her child's school, for example) her options fall into three general categories. One option is to express her dissatisfaction to the institution (for shorthand, the *voice option*). She can, for example, register a complaint with her grocery store's manager or vote in a school board election. A second option is *exit*. She can choose another grocery store or switch her children to another school. In a market where consumers have choice, exit can cause loss of revenue, which would, in turn, create the incentive for the institution to improve quality and avert this sort of loss. A formal system of public school accountability potentially expedites these processes by increasing the availability of information which the public may use in making voting or enrollment decisions. The exit option has received a modest degree of attention in work on public institutions both in economics and political science. In this context the option has typically been cast as a decision to leave one public jurisdiction for another in order to shop for a particular bundle of taxes and public services. Public institutions attempt to retain and attract constituents so as to maintain their revenue stream.[1] These voice and exit strategies can foster responsiveness on the part of the institution with which the consumer or citizen is dissatisfied, and, thus, offer potential mechanisms for indirect accountability. But the logic depends on a consumer's ability to exit, that is, to exercise choice.

While research evaluating the effectiveness of school accountability systems has increased in tandem with the spread of these systems, it has so far tended to focus

[1] See Tiebout (1956) for the classic statement on exit in the context of public institutions.

only on the impact of institutional policy prescriptions built into the program design. It has almost entirely left out the other category of consequences - public responses to the release of performance information. Yet, this speaks directly to education policy literature on choice and accountability. This paper evaluates the effect of school performance information on student exit, i.e. on the decision to exit relatively worse performing schools.

In general, studies of student mobility have focused either on patterns of movement (within or between school districts) or on consequences of movement for student well-being and academic progress (e.g. see Alexander et al 1996, and Hanushek et al 2003). Kerbow (1996) attempts to extend this discussion by investigating the extent to which school characteristics influence exit decisions. But even in this study, school indicators are subjective measures not directly linked to data on performance (e.g., survey responses about student confrontations with teachers or whether a school's curriculum seemed well-suited for the child's needs). This neglect is most likely due to the lack of extensive data on school quality before the advent of accountability systems that necessitate the collection of these data. So, while the literature on mobility has been silent as to the role of school quality in exit decisions, the emergence of new data permits an initial examination here.

My analysis uses administrative data on the population of Florida's public school students. Estimation of causal effects follow a regression discontinuity design to exploit a feature of the state's accountability system by which schools receive categorical grades based upon a continuous performance index. The remainder of my paper proceeds as follows. Section two summarizes the public school accountability system in Florida. Section three details the method for estimating the effects of school evaluations on student mobility. Sections four and five presents main results drawn from each of the four cutoffs between categorical performance grades while section

six presents results from an analysis of a subset of schools around the D-F cutoff for which the Florida accountability institutionalizes mechanisms to support student mobility out of poorly performing schools. Section seven looks at student sorting from another angle - the impact of school performance evaluations on enrollment increases. Finally, section eight concludes.

II. School Accountability in Florida

Before proceeding, it is worth considering the structure of Florida's public school accountability and choice systems. School performance evaluations take the form of a letter grade to denote quality: A, B, C, D, or F. The current system of accountability has been in place since 2002. Under this system, school grades are based on a 0 to 600 continuous grade point index. The grade points a particular school receives are based on the level of its students' achievement on the FCAT tests as well as on the amount of growth (i.e. improvement) between the previous and current year achieved by the school's students on these tests. A school's grade is then assigned according to the interval on the grade point continuum where its grade point value lies.[2] Schools with 410 or more points receive an A; schools with 380 to 409 points receive a B; schools with 320 to 379 points receive a C; schools with 280 to 319 points receive a D; and schools with fewer than 280 points receive a F.

These grades are linked to formal consequences - particularly in the case of a school receiving an F grade twice in a four-year period. Among these consequences are two directly relevant to student exit. First, students who are assigned to attend the school in the academic year immediately following receipt of the second F or who

[2]There are a few additional requirements that can force a school into a lower grade category than its grade points would indicate. These requirements have to do with the percent of students tested, adequate progress among the lowest performing students, and similarity of learning grades in reading between the lowest performing students and all eligible students. In the adjacent grade comparisons made below, those few schools which could not have reached the higher grade regardless of points due to these additional requirements are excluded from the analysis.

attended the school in the academic year for which the second F was assigned are permitted to switch enrollment to another Florida public school with a grade of C or higher. The local school district must offer at least one school within the district to which eligible students may transfer. The district maintains responsibility for the entire cost of transportation for any students who take up the option to transfer to another school in the district. Students may also transfer to a public school outside the school district if classroom space permits; however, in that case, the student's family is responsible for the cost of transportation. Second, during the period of my analysis, the students in these schools receiving an F grade twice in four years were also offered a state-funded voucher to attend private school.

Thus, for the vast majority of public school students who attend all but the worst performing schools, there are no formal institutional mechanisms for choice. For these students, residence determines school assignment. If parents wish to exercise the exit option, they must be able to relocate to another neighborhood or district. Formal mechanism of choice exist only for those in in the worst performing schools. Their parents can choose new schools - public or private - without changing residence.

Florida's public school accountability system offers an excellent opportunity to investigate the impact of school performance reports on the behavior of parents and students for at least two reasons. First, taking the form of the familiar A through F grade scale, reports of school performance in Florida are readily interpretable by the public. Even if most Floridians remain unaware of the calculations by which these grades are determined, they likely have intuitive knowledge about the rank ordering of the grade categories - that a grade of B is better than a grade of C, or that an F grade is the worst category, for example. Compare this to the federal NCLB requirement that states also classify all public schools as meeting or failing to meet Adequate Yearly Progress (AYP). It is unlikely that more than a small share of the public at

large have a sense of what AYP means or how exactly to interpret AYP assessment as an overall evaluation of school quality relative to other schools. In other words, if the report card system of school accountability - i.e. the process by which parents and citizens indirectly hold schools and school officials accountable through exit or voice - can be expected to operate effectively, it should be expected to happen in Florida where the public is provided with an easily interpreted and easily obtained summary of school quality information.

Second, and more importantly, the assignment of grades based on an underlying numeric continuum with hard cutoffs between grade categories lends itself to a regression discontinuity (RD) design for causal analysis. The RD approach permits an approximation of random assignment of treatment status (in this case the lower grade) and control status (in this case the higher grade) to otherwise comparable groups on either side of each cutoff. The details of this application are provided in the next section.

III. Method for Estimating the Effect of School Grade on Student Mobility

A straightforward, simple comparison of mobility among students in high and low performing schools would fail to provide valid estimates of the effect of school grades for at least two reasons. First, such an approach compares different kinds of students in different kinds of schools and likely suffers from selection bias. It might be that students in schools with lower school grades have a greater propensity toward mobility irrespective of the grade their schools receive. Or, it might be that student mobility and school quality (as reflected in the school grade) are both driven by other factors that distinguish students in high and low performing schools, e.g. family structures, poverty, crime, neighborhood effects, etc. In either case, a standard regression approach yields biased estimates. Second, an ideal (and seemingly counterintuitive)

approach would separate performance labels such as school grades from actual school performance. That is, the data would permit isolation of a public designation of quality apart from actual school quality (i.e. the effect of the information about school quality rather than the effect of the school quality itself). Estimates from a simple standard regression approach that enters school grades as covariates would be due to school performance generally rather than to the provision of information about school quality. A better approach would compare mobility among students in schools with equivalent or near-equivalent levels of performance but different performance labels.

In the absence of random assignment of treatment, the RD approach provides a far better alternative for estimating causal effects of school grades than a standard regression approach. Education policy researchers have increasingly made use of this approach that approximates the random assignment of field trials by comparing subjects that lay on either side of an arbitrary cutoff line (Chiang 2008; Jacob and Lefgren 2004; Ludwig and Miller 2007). The subjects placed in the control and treatment groups adjacent to the cutoff are potentially comparable. The four thresholds in the continuum of school grade points that divide the school grade categories are analyzed separately. At each threshold between school grades, treatment is attendance at a school receiving the lower grade. Thus, the analysis focuses its comparison on only those students in schools earning grade point values slightly above and slightly below these thresholds. Because the groups and their schools on either side of each threshold show similar characteristics in one year, the unobservable factors influencing their behavior the following year are assumed to have no difference on average in the next year. Therefore, any actual differences in mobility patterns in the second year are attributed to the impact of attending a school assigned a lower school grade.[3]

I focus on the first set of school grades released under the new calculation rules. Academic years 2001-2002 and 2002-2003 are used. Analyzing this time frame is

[3]To be sure, students exit schools for many reasons. These reasons probably frequently have nothing to do with school quality (e.g. families move after parents change jobs). These data do not identify the reasons for each child's exit. However, this limitation is not terribly consequential. The key argument in the regression discontinuity is that the schools on either side of the cutoff point between school grades are balanced on unobserved characteristics. In other words, the proportion of exits due to reasons other than the school performance grade is unlikely to be much different among schools who score just above and just below the cutoff between an A and B grade for example. Therefore, the differences across the schools can be attributed to the grade.

particularly advantageous. It is unlikely that schools anticipated their school grades at the end of the 2001-2002 school year. Despite the fact that the new grading system was made public in late 2001, the new rules were complex enough to exclude this possibility (Chiang 2008). Thus, this set of grades serves as an "exogenous shock" on the schools and their student populations (West and Peterson 2006).

I let y_{is} be a binary indicator indicating exit for student i from school s, where s is an index of school attended in year 0. Enrollment is observed in the spring of each year as the school where students take the FCAT. Exit is defined as enrollment in a school during year 2002-2003 different from that during year 2001-2002. This can be observed in two ways: 1) Students whose 2002-2003 record indicates a public school other than that attended in the previous year have changed from one Florida public school to another public school in the state; and 2) Students who have a record indicating enrollment in a Florida public school in 2001-2002 but have no record for the 2002-2003 have exited Florida public schools entirely, transferring either to a public school in another state or to a private school. Two types of moves, however, cannot be considered voluntary exits from year 2001-2002 schools. Students who attended year 2001-2002 schools that closed at the end of that school year are excluded from the analysis. Similarly, students who moved for structural reasons, that is, whose grade level in year 2002-2003 exceeds the highest grade level served by their year 2001-2002 school (e.g. a student who attends a year 2001-2002 school that only serves the first through fifth grades and who enters the sixth grade in year 2002-2003) are excluded from the analysis. I also let $GrdPts_s$ be the grade point value earned by school s at the end of year 2001-2002 and G_s be a binary indicator for whether school s received the lower of the two grades surrounding each threshold (e.g. receipt of a B relative to an A at the A-B threshold; receipt of a C relative to a B at the B-C threshold; receipt of a D relative to a C at the C-D threshold; and receipt of an F relative to a

D at the D-F threshold) at the end of year 2001-2002.

For each threshold, I estimate the expected likelihood of exit among students attending a year 2001-2002 school with a lower school grade (relative to the higher grade on the other side of the threshold) and a school grade point value exactly at the threshold value. This quantity is estimated using lower graded schools with grade point values within a specified distance from the threshold; this distance is the bandwidth, h. The equation is:

$$y_{is} = logit^{-1}[\alpha_{Lower} + \beta_{Lower}(GrdPts_s - Threshold) + \varepsilon_{is}],$$

such that $-h < GrdPts_s - Threshold < 0$, and where $Threshold$ equals the school grade point value marking the cutoff between the upper and lower school grade.

Similarly, for each threshold, I also estimate the expected likelihood of exit among students attending a year 0 school with a higher school grade (relative to the lower grade on the other side of the threshold) and a school grade point value exactly at the threshold value. This quantity is estimated using higher graded schools with grade point values within the same bandwidth as the previous equation. The equation is:

$$y_{is} = logit^{-1}[\alpha_{Higher} + \beta_{Higher}(GrdPts_s - Threshold) + \varepsilon_{is}],$$

such that $0 < GrdPts_s - Threshold < h$. This local regression procedure follows the sharp regression discontinuity design presented by Imbens and Lemieux (2007) and implemented in a similar context by Chiang (2008), but adapted to handle a binary outcome.[4] [5]To implement the design, at each threshold I combine these local estimations for students in above and below threshold schools into a single regression and include X, a vector of covariates:

[4]In the "Sharp" regression discontinuity context the treatment variable is a deterministic function of one of the covariates. In other words, the probability of receiving treatment changes from zero to one at some threshold level of the forcing covariate, as opposed to a fuzzy regression discontinuity context where the probability of assignment to treatment jumps some smaller degree at the threshold.

[5]Chiang (2008) uses a RD design applied to Florida school grades to estimate the effect of a school's receipt of an F grade on student test score performance the following year.

$$y_{is} = logit^{-1}[\beta_0 + \beta_1 G_s + \beta_2(GrdPts_s - Threshold) + \beta_3(G_s \times (GrdPts_s - Threshold)) + \gamma X_{is} + \varepsilon_{is}].$$

In this equation, the estimated coefficient $\widehat{\beta}_1$ on the indicator of the lower school grade captures the causal relationship between school grade and likelihood of exit. Covariates include controls for student demographics: and indicator for black, an indicator for Hispanic, an indicator for eligibility in free/reduced priced lunch program, an indicator for female, and year 2001-2002 test score on the FCAT reading and math tests. The scale test scores are standardized by subject and grade level. I exclude certain students not subject to the state's accountability program. These include students with one of several disabilities and students enrolled in English for Speakers of Other Languages.[6]

The student-level data used in this paper come from administrative data on the population of Florida's public school students provided by the Florida Department of Education. The data include student demographic characteristics, grade level, FCAT reading and math tests scores, special education and English for Speakers of Other Languages classification, school district and school attended for third through tenth grade public school students in the 2001-2002 and 2002-2003 academic years. School level data include school grade points and school grades for the 1998-1999 through 2002-2003 academic years.

Estimates from RD designs are potentially highly sensitive to bandwidth selection - i.e. the distance from the threshold on the continuum of the forcing variable used to determine which cases are included in the regression. Smaller bandwidths can give greater weight to unusual groups of the most improving schools in the lower grade and the most lagging schools in the higher grade on the opposite side of the threshold. Smaller bandwidths also mean for fewer schools in the comparisons; although thousands of students may remain in the subsample used for estimation, it must be remembered that these students are clustered within a smaller number of schools.

[6] All analyses were repeated including these students. In no case were the results substantively different.

On the other hand, a larger bandwidth can include cases that are further apart along the continuum of the forcing variable and, therefore, no longer comparable on the covariates that might confound treatment effects. Unfortunately, there is no universally accepted standard for bandwidth selection. However, one popular approach is to select a bandwidth that minimizes a cross-validation criterion (Ludwig and Miller 2005; Imbens and Lemieux 2007; Chiang 2008). Here I follow the procedure as described by Chiang. The aim is to minimize the expected squared prediction error at the threshold. To estimate this quantity for a given threshold, I begin with two groups of schools. One group G_L includes the 50% of schools closest to the threshold on its lower side based on 2001-2002 school grade points; the other group G_H includes the 50% of schools closest to the threshold on its higher side. For each school k within the lower group, I separately regress schools' proportion of leavers on year 2001-2002 grade points using only schools in that group with a specified bandwidth h to the left of that school on the grade point scale. Using the results from the regression, I obtain the predicted proportion of leavers $\hat{\overline{y}}_{k,h}$for school k at this bandwidth. After repeating the regression to the left of each school, the cross validation criterion for the specified bandwidth is constructed as the average mean squared error: $CV_{G_L,h} = \frac{1}{G_L}\sum_{k \in G_L}(\overline{y}_k - \hat{\overline{y}}_{k,h})^2$. Then, construction of the cross-validation criterion is repeated for each specified bandwidth from 0 to 40. The bandwidth that comes closest to minimizing the cross-validation criterion is prefered. This is the smallest bandwidth value that should be used. The entire process is repeated for the group of schools above the threshold in the same way except that regressions use schools to the right of school k. The appendix contain plots of the cross-validation criterion for each group at each bandwidth. At the A-B threshold, bandwidths of 13 and 20 minimize the criterion for A and B schools respectively. At the B-C threshold, bandwidths of 28 and 34 minimize the criterion for B and C schools respectively. At the

C-D threshold, bandwidths of 20 and 26 minimize the criterion for C and D schools respectively. At the D-F threshold, bandwidths of 25 and 28 minimize the criterion for D and F schools respectively. While it is legitimate to use different bandwidths on each side of a given threshold, for ease of presentation I select a common bandwidth for both sides of each threshold; this is the greater of the two bandwidth values that minimize the cross validation criterion on each side. These are the bandwidth used for the results presented in this paper. As an added check, analyses are duplicated across several bandwidths of five, ten, twenty, and thirty grade points. The results for these additional bandwidths appear in the appendix.

Using the preferred bandwidths describes above, table 1 shows the mean level of covariates among students in school within the bandwidths in year 2001-2002. The table also includes the predicted values for these covariates at the threshold estimated from local linear regressions on each side. In a discontinuity design, a validity check is to see if the predicted values of the covariates differ significantly at the threshold depending on the side from which they are estimated. Table 1 shows that in almost all cases, even when the samples on each side of a threshold are not perfectly balanced, the differences are not significant at the 5% level. Nevertheless, due to imperfect balance covariates are controlled in the regressions.

Before considering estimates from the regressions using student-level data, I present graphical depictions of the effect of lower school grades on student mobility at the school-level. Figures 1 through 4 plot year 2001-2002 school grade points against the proportion of students who exit between years 2001-2002 and 2002-2003 for each pair of adjacent school grade categories. The figures also include lines displaying local linear regressions of proportion of students who exit on school grade points. Ignoring the school grades for a moment, a reasonable expectation would be that downward sloping regression lines indicating fewer moves from schools of higher quality. Reintroducing the grade categories, a reasonable expectation would be downward sloping

Table 1: Covariate Mean and Predicted Values at Each Threshold

			Black	Hisp.	Female	Free/Red.	Read	Math
A-B	Mean	A	.175	.125	.514	.320	.347	.354
h=20		B	.175	.143	.519	.407	.242	.265
	Pred.	A	.202	.151	.516	.352	.316	.310
		B	.151	.193	.512	.398	.277	.296
B-C	Mean	B	.198	.157	.517	.431	.218	.235
h=30		C	.291	.170	.517	.519	.057	.088
	Pred.	B	.254	.159	.517	.492	.149	.157
		C	.244	.163	.517	.474*	.139	.163
C-D	Mean	C	.401	.164	.518	.543	-.084	-.050
h=26		D	.600	.211	.520	.720	-.400	-.312
	Pred.	C	.489	.160	.512	.580	-.179	-.159
		D	.559	.225	.531*	.645	-.290*	-.207
D-F	Mean	D	.574	.108	.516	.659	-.332	-.251
h=28		F	.818	.100	.529	.666	-.524	-.428
	Pred.	D	.721	.116	.503	.732	-.505	-.394
		F	.813	.092	.533	.589	-.429	-.325

Notes: Means are mean values for schools within the specified bandwidth. Predicted values are values of covariates at threshold predicted from local linear regressions using the specified bandwidths. Reading and math test scores were standardized by student grade level. Asterisks indicate predicted values among with the lower grade differ significantly at the 5% level from those with the higher grade.

regression lines on both sides of the cutoff between grades but with a drop at the threshold. In other words, the regression line for the lower graded schools would end at the threshold with a greater proportion of leavers while the higher graded schools' regression line would begin at the threshold with a smaller proportion of leavers. This is what appears for the A-B cutoff in figure one with the preferred bandwidth. On the other hand, the expected drop in rate of exit fails to appear at the other thresholds examined using the preferred bandwidths.

Figure 1: Proportion Exit by School Grade (A-B)

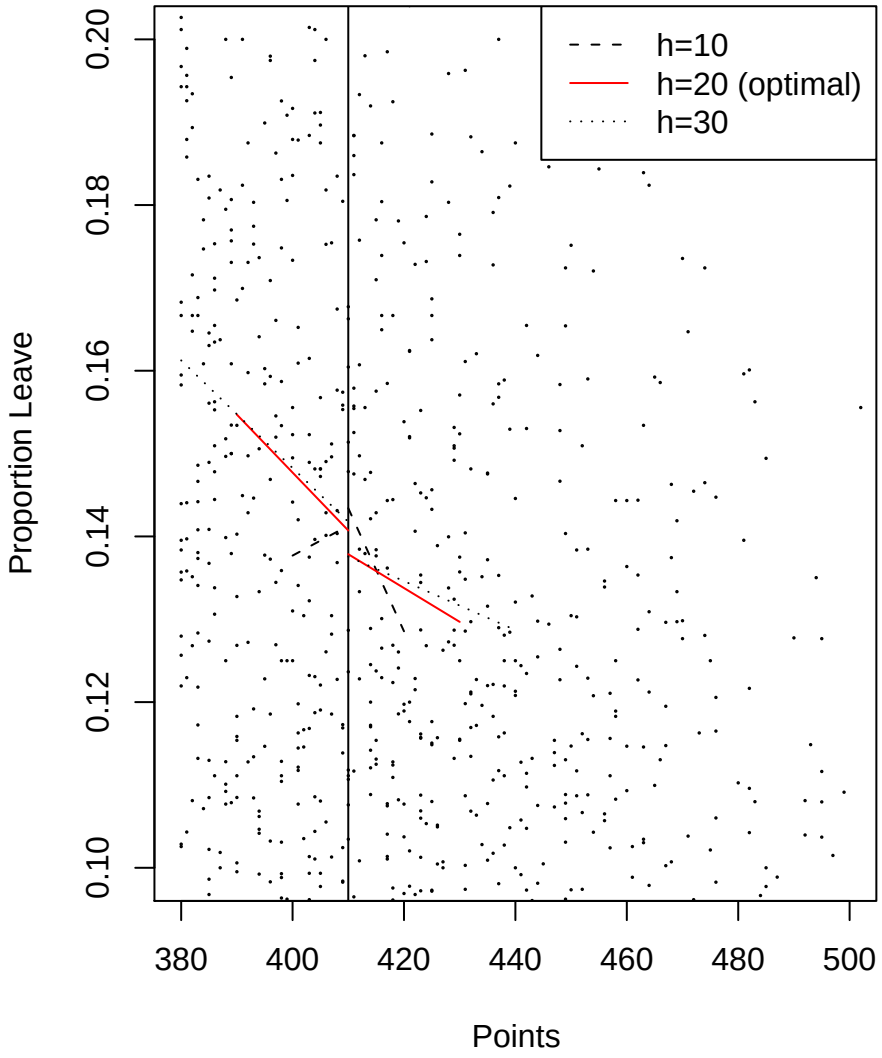


Figure 2: Proportion Exit by School Grade (B-C)

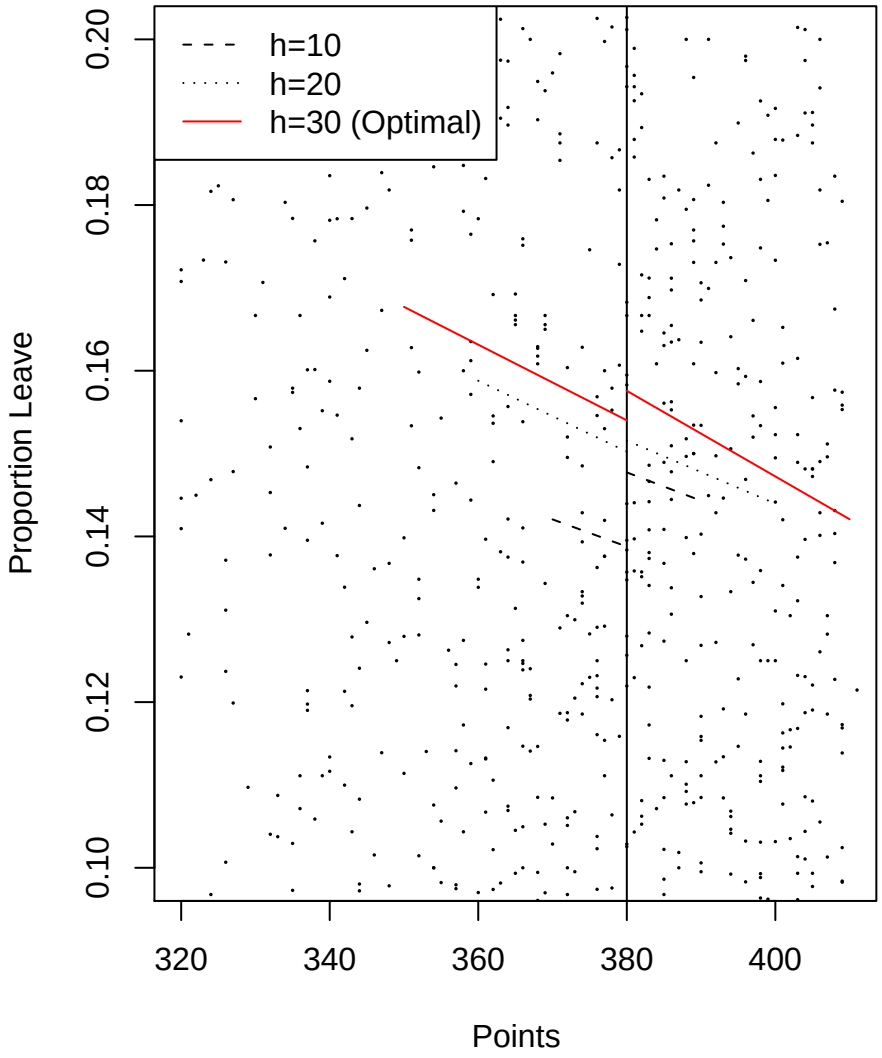


Figure 3: Proportion Exit by School Grade (C-D)

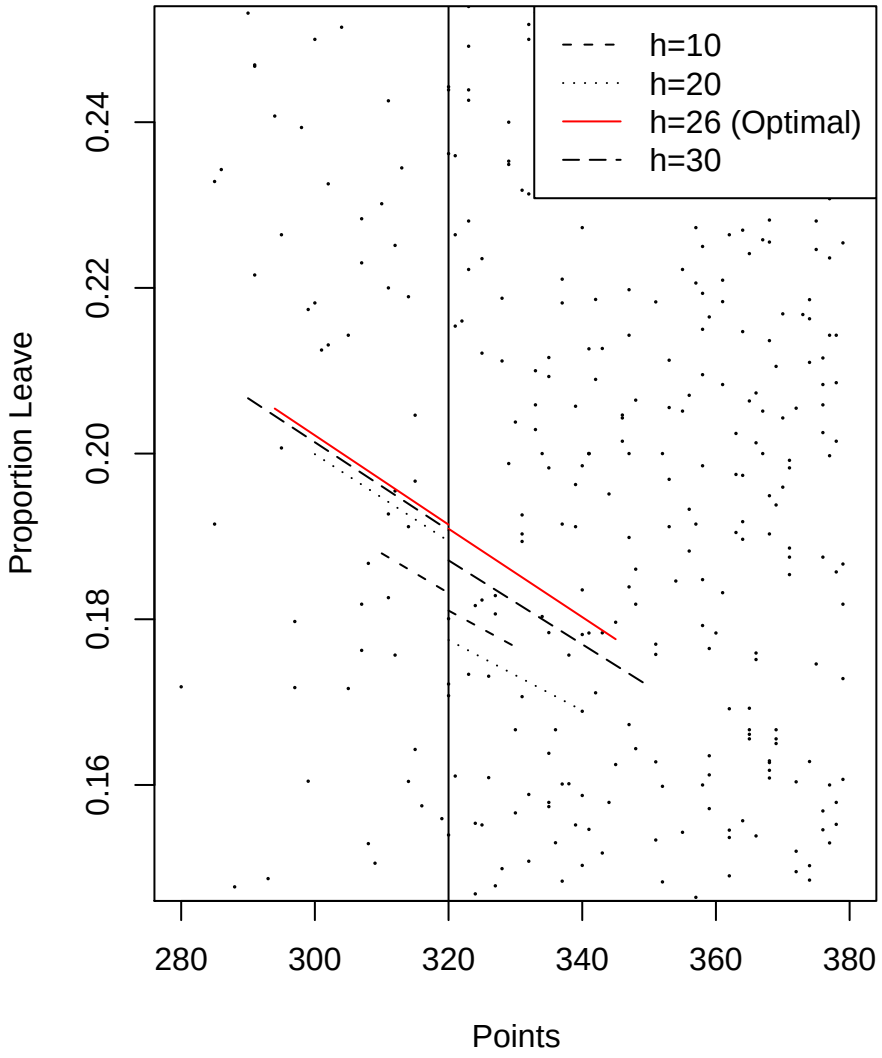
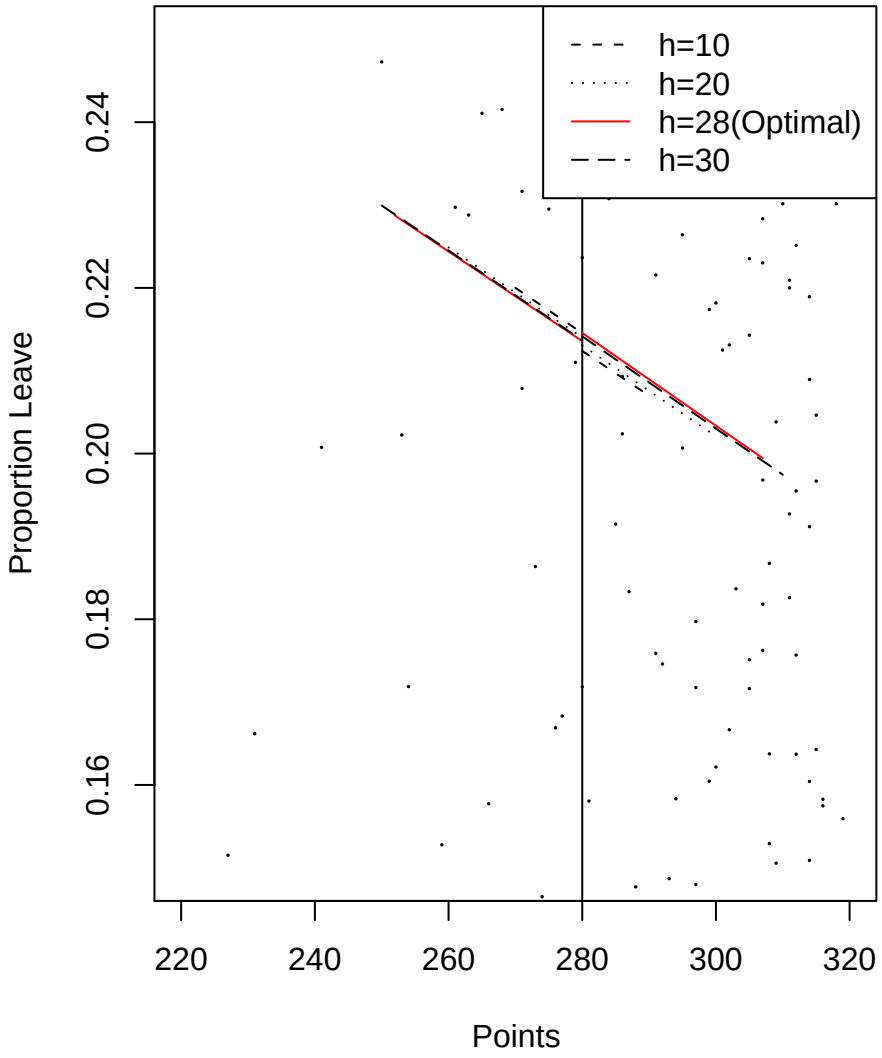


Figure 4: Proportion Exit by School Grade (D-F)



IV. Estimation Results

Figures 1 through 4 present mixed evidence for an exit effect of school grades at the aggregate level, varying across the set of cutoffs between grade categories. To test a formal hypothesis about the effect on school grades on student exit at the

Table 2: Regression Coefficients at A-B Threshold with 20 Point Bandwidth

	Point Est.	Std. Err.	First Diff.	Std. Err.
(Intercept)	-2.08	0.110	-	-
B	0.071	0.185	0.009	0.173
School Grd Pts	0.005	0.010	-	-
School Grd Pts * B	-0.005	0.014	-	-
Black	0.203	0.038	0.020	0.005
Hispanic	0.118	0.073	0.012	0.008
Female	-0.066	0.015	-0.006	0.001
Free/Red Lunch	0.434	0.064	0.046	0.007
FCAT Read	-0.231	0.027	-0.020	0.003
FCAT Math	-0.119	0.013	-0.011	0.001

Notes: Point estimates are from logit regression. Standard errors are clustered by school. First differences reflect the increase in probability of an exit for a one unit change in the covariate while holding other covariates at the following specified values: school grade at zero (i.e. the higher grade); school grade points at the threshold value; black, Hispanic, female, and free/reduced price lunch at zero; and standardized FCAT scale scores at their mean values. For school grade and demographics, the first difference captures the move from a zero to a one on these binary indicators. For test scores the first difference captures a one standard deviation increase in scale test score. First differences are estimated through simulation.

individual level I estimate the equation presented in the last section for each of the four thresholds separately. In each case multiple covariates are present, which appear in table 1.

Tables 2 to 5 present the estimates of school grade effects on student exit at each of the thresholds. The estimates of interest in these tables are those for the grades: B, C, D, and F respectively. In each case, these estimates capture the effect of receiving a lower school grade (relative to the next highest grade) on student exit. The first difference for these grades captures the difference in the probability of exit between receiving the lower grade rather than the higher grade estimated at the cutoff between grades. At each of these cutoffs, this estimate is too small and estimated too imprecisely to be statistically distinguishable from zero. Thus, estimates from each of the thresholds do not indicate that publicly released school quality reports cause parents to sort their children out of lower graded schools. In turn, these results cast doubt on the notion that report card systems foster the exit option as an indirect mechanism of school accountability.

V. Subgroup Analysis

Table 3: Regression Coefficients at B-C Threshold with 30 Point Bandwidth

	Point Est.	Std. Err.	First Diff.	Std. Err.
(Intercept)	-2.013	0.054	-	-
C	0.149	0.095	0.016	0.010
School Grd Pts	0.002	0.007	-	-
School Grd Pts * C	0.001	0.007	-	-
Black	0.178	0.033	0.019	0.004
Hispanic	-0.032	0.053	-0.003	0.005
Female	-0.070	0.012	-0.007	0.001
Free/Red Lunch	0.418	0.043	0.049	0.006
FCAT Read	-0.197	0.018	-0.019	0.002
FCAT Math	-0.116	0.011	-0.011	0.001

Notes: Point estimates are from logit regression. Standard errors are clustered by school. First differences reflect the increase in probability of an exit for a one unit change in the covariate while holding other covariates at the following specified values: school grade at zero (i.e. the higher grade); school grade points at the threshold value; black, Hispanic, female, and free/reduced price lunch at zero; and standardized FCAT scale scores at their mean values. For school grade and demographics, the first difference captures the move from a zero to a one on these binary indicators. For test scores the first difference captures a one standard deviation increase in scale test score. First differences are estimated through simulation.

Table 4: Regression Coefficients at C-D Threshold with 26 Point Bandwidth

	Point Est.	Std. Err.	First Diff.	Std. Err.
(Intercept)	-1.809	.089	-	-
D	-0.064	0.121	-0.008	0.015
School Grd Pts	0.002	0.005	-	-
School Grd Pts * D	-0.009	0.009	-	-
Black	0.112	0.056	0.014	0.007
Hispanic	-0.220	0.064	-0.025	0.008
Female	-0.074	0.018	-0.009	0.002
Free/Red Lunch	0.311	0.049	0.043	0.006
FCAT Read	-0.174	0.018	-0.020	0.002
FCAT Math	-0.101	0.017	-0.012	0.002

Notes: Point estimates are from logit regression. Standard errors are clustered by school. First differences reflect the increase in probability of an exit for a one unit change in the covariate while holding other covariates at the following specified values: school grade at zero (i.e. the higher grade); school grade points at the threshold value; black, Hispanic, female, and free/reduced price lunch at zero; and standardized FCAT scale scores at their mean values. For school grade and demographics, the first difference captures the move from a zero to a one on these binary indicators. For test scores the first difference captures a one standard deviation increase in scale test score. First differences are estimated through simulation.

Table 5: Regression Coefficients at D-F Threshold with 28 Point Bandwidth

	Point Est.	Std. Err.	First Diff.	Std. Err.
(Intercept)	-1.453	0.122	-	-
F	0.134	0.176	0.024	0.031
School Grd Pts	0.007	0.007	-	-
School Grd Pts * F	-0.019	0.011	-	-
Black	-0.220	0.118	-0.034	0.018
Hispanic	-0.600	0.135	-0.081	0.019
Female	-0.045	0.026	-0.007	0.004
Free/Red Lunch	0.291	0.054	0.052	0.009
FCAT Read	-0.163	0.024	-0.024	0.004
FCAT Math	-0.098	0.019	-0.015	0.003

Notes: Point estimates are from logit regression. Standard errors are clustered by school. First differences reflect the increase in probability of an exit for a one unit change in the covariate while holding other covariates at the following specified values: school grade at zero (i.e. the higher grade); school grade points at the threshold value; black, Hispanic, female, and free/reduced price lunch at zero; and standardized FCAT scale scores at their mean values. For school grade and demographics, the first difference captures the move from a zero to a one on these binary indicators. For test scores the first difference captures a one standard deviation increase in scale test score. First differences are estimated through simulation.

Perhaps the results from the previous section mask effects that occur only among specific subgroups. To examine this possibility, I repeat the analyses from the previous section applied only to subsamples of students based on test performance, race/ethnicity, free/reduced lunch status. These results are presented in table 6. The table includes first differences in the probability of exit for that subgroup when receiving a lower school grade. The results confirm those of the last section. In no case do the estimates confirm that receiving a lower school grade prompts student exit in the following academic year.

VI. Institutional Mechanisms for Choice

The results in the previous section certainly contradict the idea that the public uses the information provided through report card systems to hold schools accountable through the exit option. However, in another sense the results are unsurprising. Public school systems tend to heavily regulate school enrollment through attendance zones and other mechanisms whereby pupils are assigned a school based on residence. These regulations limit the potential for the exit option by effectively increasing the

Table 6: Difference in Probability of Exit by Subgroup

	A-B Threshold		B-C Threshold		C-D Threshold		D-F Threshold	
	First Diff.	Std. Err.	First Diff.	Std. Err.	First Diff.	Std. Err.	First Diff.	Std. Err.
Read: Level 1	-0.009	0.016	0.019	0.013	-0.004	0.018	0.040	0.033
Read: Level 5	0.071	0.056	0.017	0.019	-0.007	0.046	-	-
Math: Level 1	0.001	0.018	0.022	0.012	-0.008	0.016	0.012	0.025
Math: Level 5	0.030	0.027	-0.006	0.013	-0.014	0.036	-	-
Black	-0.016	0.020	0.028	0.016	0.002	0.020	0.013	0.028
Hispanic	0.013	0.045	0.029	0.020	-0.021	0.023	0.055	0.049
White	0.014	0.018	0.007	0.012	-0.019	0.021	0.030	0.056
Free/Reduced	-0.011	0.015	0.022	0.014	-0.015	0.023	0.026	0.039
Not Free/Red.	0.020	0.028	0.015	0.013	0.000	0.018	0.024	0.036

Notes: First differences reflect the increase in probability of an exit for a one unit change in the covariate while holding other covariates at the following specified values: school grade at zero (i.e. the higher grade); school grade points at the threshold value; black, Hispanic, female, and free/reduced price lunch at zero; and standardized FCAT scale scores at their mean values. For school grade and demographics, the first difference captures the move from a zero to a one on these binary indicators. For test scores the first difference captures a one standard deviation increase in scale test score. First differences are based on estimated coefficients from logit models using the bandwidths noted above.

cost of such an action. For a parent to pull her children out of a school that received an F, she must have another place in which to enroll them. Typically there are only four options for this parent who is unhappy with the performance of her children's school: 1) move to another residence where her children will be assigned to another public school; 2) enroll her children in private or parochial school; 3) pull her children out of all professional schooling and home-school them; or 4) simply remain in the assigned public school despite its performance evaluation. Each of these options - save the fourth - is costly in the short term. Parents may lack the resources, financial or otherwise, to engage in any of the first three options. Instead, we may expect them to take the fourth option - the only one which does not entail exiting the original school. If this is the case then it hardly seems possible for these accountability systems to foster the indirect accountability of the exit option simply by providing information about school quality. It would seem then that indirect accountability depends on the capacity for choice, which most public school systems do not provide institutional mechanisms to support.

However, two institutionalized mechanisms to reduce the cost of exit were among the consequences linked to school grades during the period analyzed in this paper. In 2002 the Florida accountability system mandated that students attending any school receiving two F grades in a four year period (1998-1999 through 2001-1002 in this case) would be offered a seat in at least one other public school with a grade of C or higher to attend, with transportation costs paid by the district, or a voucher to attend a private school.[7] For these parents and students, the accountability system not only provides quality information about schools but also offers a mechanism for transfer that would reduce the cost of exit. Therefore, by restricting the sample to students in just those schools that received one F prior to the 2001-2002 academic year and examining those surrounding the D-F cutoff in 2002, a stronger effect of of school grade might manifest.

Only 21 schools received a grade of D at the end of the 2001-2002 school year after receiving a grade of F in a prior year, enrolling 3,512 students.[8] Only nine schools received their second grade of F at the end of this year, enrolling 1,766 students. As shown below, analysis of the 1,766 students in these schools yields a reasonable degree of evidence about accountability and exit. Results from a RD analysis at varying bandwidths appear in table 7. The same controls are included as in the previous section.

Only one school falls within the range of a ten point bandwidth on the F-side of the threshold, so estimates are not possible there. The twenty and thirty point bandwidths allow inclusion of more schools without deviating too far from the threshold. Although less than an ideal number of schools fall into these groups of F schools, the results are striking. Receipt of an F school grade increase the likelihood of exit

[7] The Florida Supreme Court has since ruled the voucher option unconstitutional.

[8] These figures include the same students as before: structural moves, students in schools closing at the end of the 2001-2002 academic year, students classified with certain disabilities that exclude them from the calculation of school grades, and students enrolled in English for Speakers of Other Languages. Regressions were repeated including these last two types of exclusions, but results remain substantively unchanged.

Table 7: Effect of 2002 School F Grade on Student Exit
for Students in Schools with a Prior F Grade

Bandwidth	Point Est.	Std. Err.	First Diff.	Std. Err.	D Schools	D Students	F Schools	F Students
All	1.236	0.297	0.282	0.072	21	3,512	9	1,766
28 (Optimal)	1.652	0.331	0.381	0.072	19	3,223	7	1,131
20	1.856	0.302	0.423	0.065	16	2,765	5	716

Notes: Estimates are from separate logit regressions. Standard errors are clustered by school. First differences reflect the increase in probability of exit of a B school grade relative to an A school grade. First differences are estimated through simulation while setting school grade points at the threshold value, student test scores at their mean value, and student demographics at their mode values. The last four columns show the number of students and schools in each school grade group based on the specified bandwidth.

among students in those schools by approximately forty percentage points.

Students who receive a school grade of F and are provided with a an institutional mechanism for the exercise of choice (either to another public or a private school) are more likely to exit than those in schools receiving a grade of D and no institutional mechanisms for choice. Even so, it remains difficult to determine what exactly the treatment is since it has two components. One, they are enrolled in a school receiving a second F grade in a four year period. Two, they receive the transfer option. The increased likelihood of exit could be due to receipt of a signal about a particularly poorly performing school (i.e. one that has received two F grades). However this seems unlikely since it assumes that the costs which constrain residential choice in the other comparisons explored above are irrelevant here. It is much more likely that the effect is due, at least in part, to the choice options.[9] Thus, it seems that the exit option (and any incentives it provides for improving school performance) depends not only on the release of school performance reports to the public but also on providing parents with institutional mechanisms for choice.

[9]The fact that something other than just the transfer option is affecting exit is suggested by the takeup rate for vouchers, which has been estimated to be about five percent of eligible students, and the takeup rate for the public school transfer option, which has been estimated at about ten percent of eligible students (Grech 2002; Peterson 2006). The model estimated with a bandwidth of 28 grade points predicts that 26.7% of students in these F schools exit; the data indicate 29.4% of students actually exit.

VII. Entrance

Even if school performance reports do not affect the decision to leave schools in the absence of choice programs, perhaps they affect the decision about what schools to enter once the decision to exit the original school is made for other reasons. Unfortunately, the current dataset does not permit individual level analysis of entrance using the RD approach.[10] Therefore, I estimate the impact of school grades on entrance at the school level. First, I subset the data to include only students who are attending schools in 2002-2003 other than those attended in 2001-2002. The data include students who switch between public schools and who switch into public schools for year 2002-2003. Two types of student movers are excluded: those who switch for structural reasons (as described in section four of this paper) and those non-repeating students in third grade in 2002-2003. This second group is excluded because the data do not include second grade records. Therefore, I cannot verify their enrollment in the prior year. This subset provide the universe of voluntary movers entering year 2002-2003. Each of these students is matched to a school for year 2002-2003. For each school, I aggregate the number of these students entering the school. For each school, I then converted the number of entrants into a proportion of prior year (2001-2002) enrollment. This step is taken simply because larger schools are more likely to receive larger numbers of students. These proportions are then used as the outcome variable in local regressions on school grade points around each of the thresholds. Once again, I determine the preferred bandwidths using the cross-validation criterion as described in section four. The plots appear in the appendix. At the A-B threshold, bandwidths of 22 and 21 come closest to minimizing the criterion for A and B schools respectively.

[10] In the analysis of exit, the treatment is the school grade assigned to the school that the individual student attends in the first year of analysis. Thus, each individual receives a single treatment. In the case of entrance there are multiple treatments - indeed treatments are the school grades for all schools to which a student could transfer. One could attempt to sidestep this issue by using the grade of only the school to which a student transfers as treatment, but that would provide a treatment only for those students who did move (which would perfectly predict movement since this is selecting upon a single value of the dependent variable) but not for those students who did not move. Thus, the data do not permit analysis of entrance at the individual level.

Table 8: School Grade Report and Student Entrance

Threshold	Coefficient	Std. Error
A-B	0.004	0.011
B-C	0.007	0.009
C-D	0.001	0.019
D-F	-0.027	0.026

At the B-C threshold, bandwidths of 13 and 20 come closest to minimizing the criterion for B and C schools respectively. At the C-D threshold, bandwidths of 25 and 27 come closest to minimizing the criterion for C and D schools respectively. At the D-F threshold, bandwidths of 24 and 28 come closest to minimizing the criterion for D and F schools respectively. Once again, the larger of the two preferred bandwidths around a threshold is used. Using standard OLS, I regress the proportion of new students on school grade, school grade points, and an interaction of the grade and points, as well as the share of black, Hispanic, female, and free/reduced lunch students among the entrants and average reading and math test scores. The primary results of interest (i.e. the point estimates of the impact of receiving a lower grade) are contained in table 8. Reports of school performance do not appear to affect sorting into schools among students who move.

VIII. Conclusion

Whereas past analyses of student mobility have typically neglected the role of school quality in the exit decision, this paper attempt to fill that void by examining the effect of accountability grades on student exit. In general parents do not appear to be sorting their children out of schools receiving lower school accountability grades. This casts doubt on the effectiveness of report card systems to foster indirect accountability through the exit option. If parents do not sort their children out of poorly performing schools, then school officials need not worry about declining enrollment and budget cuts resulting from poor performance. This does not mean that parents ignore these performance evaluations entirely. It may be the case that parents decide to leave one school without regard to quality as measured by school grades

(e.g. residential relocation, peer group change, etc.) but then use the grades to make decisions about which schools to enroll their children in after the initial decision to make a move is made. But even in that case, when school grades are used to make the marginal decision about which school to enter once the decision to move has been made rather than to make a decision to move, the exit option as traditionally defined as a mechanism of institutional control remains in limited use.[11]

The one exception is for students who have a formal choice option not based strictly on residence. In that case information causes exit from lower performing schools. However, since most accountability programs feature only limited (if any) choice mechanisms - i.e., they typically provide families a choice among only one or two public schools which typically perform only slightly better than the original school - the prospects for school accountability systems to improve the quality of education in the United States seem unsure at best. Yet, accountability systems are among the most popular and most relied upon education reform tools today.

[11] It may also be the case that parents simply unaware of school quality. While school report cards are released directly to parents and receive substantial attention in local media outlets (especially in the year examined here when the program had recently been revised), the data do not include information about parents' knowledge of school grades. But it must be noted that this is an issue of a causal mechanism (i.e., how a cause has an effect) rather than a causal impact (i.e., whether a cause has an effect). Lack of awareness among parents would not refute these findings. Indeed, lack of awareness would only bolster the argument made here about the ineffectiveness of school reports cards for promoting indirect accountability. Under conditions that should favor awareness, public provision of information about school quality fails to foster this type of accountability.

Works Cited

Alexander, K. L., D. R. Entwisle, and S. L. Dauber. 1996. "Children in Motion: School Transfers and Elementary School Performance." *Journal of Educational Research* 90(1): 3-12.

Chiang, Hanley. 2008. "How Accountability Pressures on Failing Schools Affects Student Achievement." Working paper. Harvard University.

Grech, D. 2002. "Dade Students Using Vouchers for First Time." *The Miami Herald*, July 2, page 1A.

Hanushek, E.A., J.F. Kain, and S.G. Rivkin. 2004. "Disruption versus Tiebout Improvement: The Costs and Benefits of Switching Schools." *Journal of Public Economics*. 88(9-10): 1721-1746.

Hirschman, A.O. 1970. *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States*. Cambridge: Harvard University Press.

Imbens, G., and T. Lemieux. 2007. "Regression Discontinuity Designs: A Guide to Practice." Working Paper 13039. Cambridge: National Bureau of Economic Research.

Jacob, B., and L. Lefgren. 2004. "Remedial Education and Student Achievement: A Regression-Discontinuity Analysis." *Review of Economics and Statistics*. 68: 226-44.

Kerbow, D. 1996. "Patterns of Urban Student Mobility and Local School Reform." *Journal of Education for Students Placed at Risk*. 1(2): 147-169.

Ludwig, J., and D. Miller. 2007. "Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design." *Quarterly Journal of Economics* 122 (1): 159-208.

Tiebout, C. 1956. "A Pure Theory of Local Expenditures" *The Journal of Political Economy* 64 (5): 416-24.

West, M.R. and P.E. Peterson. 2006. "The Efficacy of Choice Threats Within School Accountability Systems: Results from Legislatively Induced Experiments." *The Economic Journal* 116 (March): C46-C62.

Appendix

Figure 5: A-B Threshold Cross-Validation Criterion for Exit

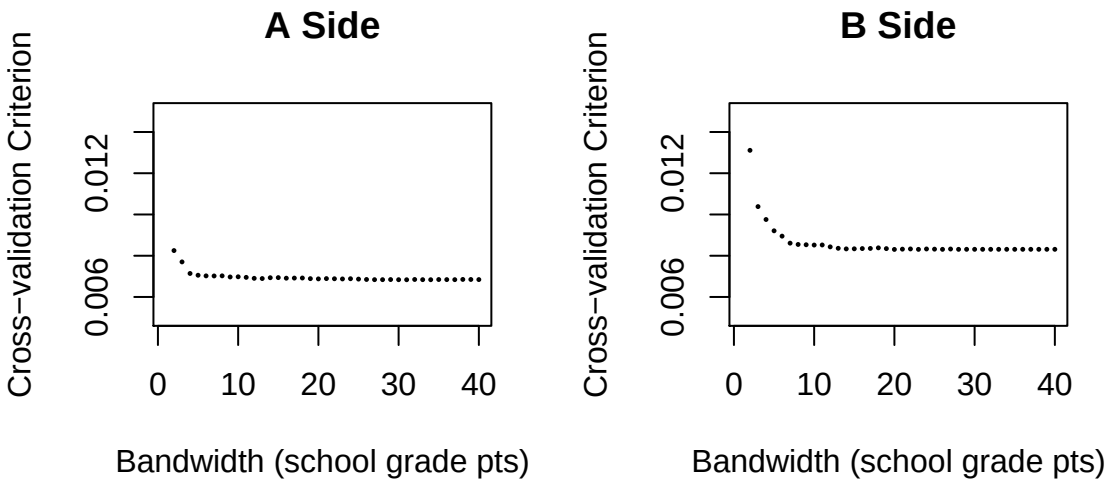


Figure 6: B-C Threshold Cross-Validation Criterion for Exit

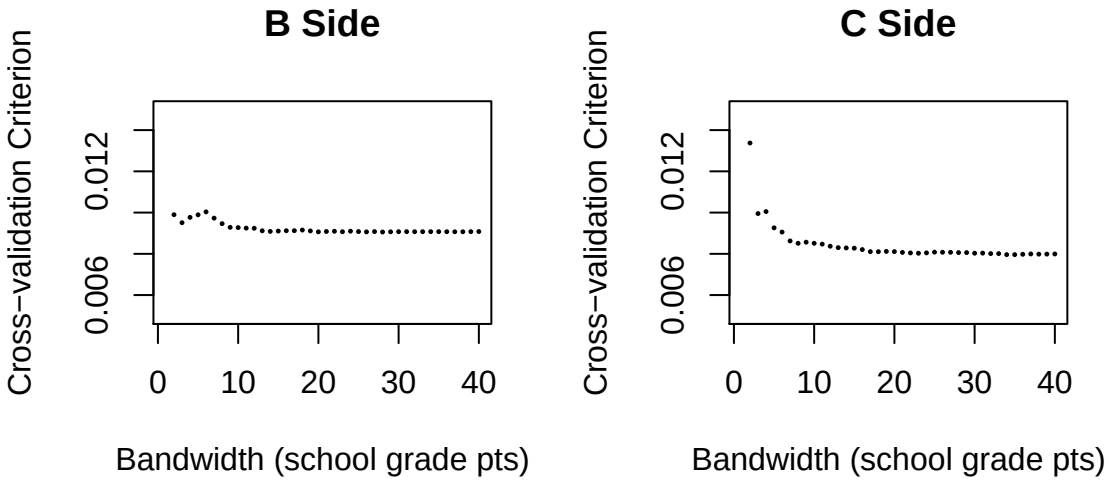


Figure 7: C-D Threshold Cross-Validation Criterion for Exit

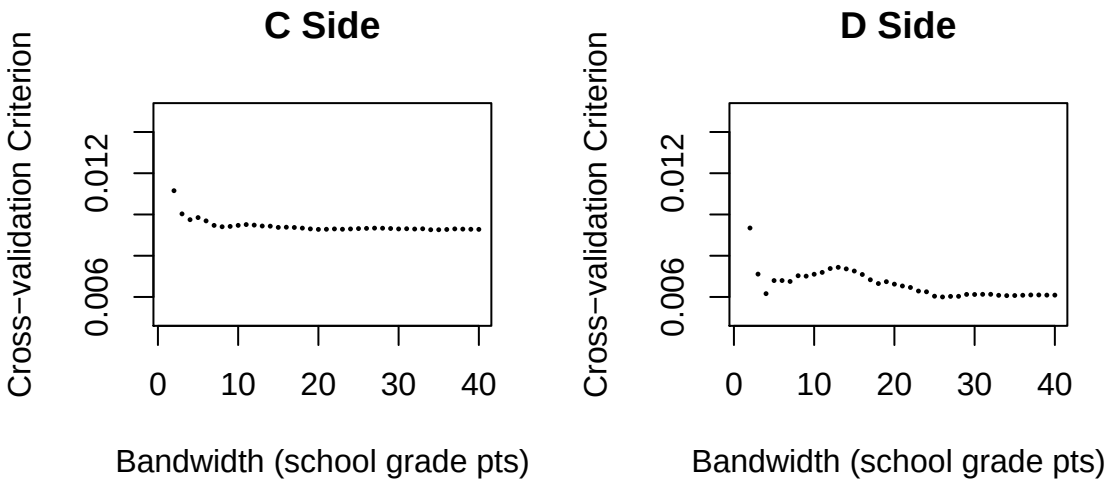


Figure 8: D-F Threshold Cross-Validation Criterion for Exit

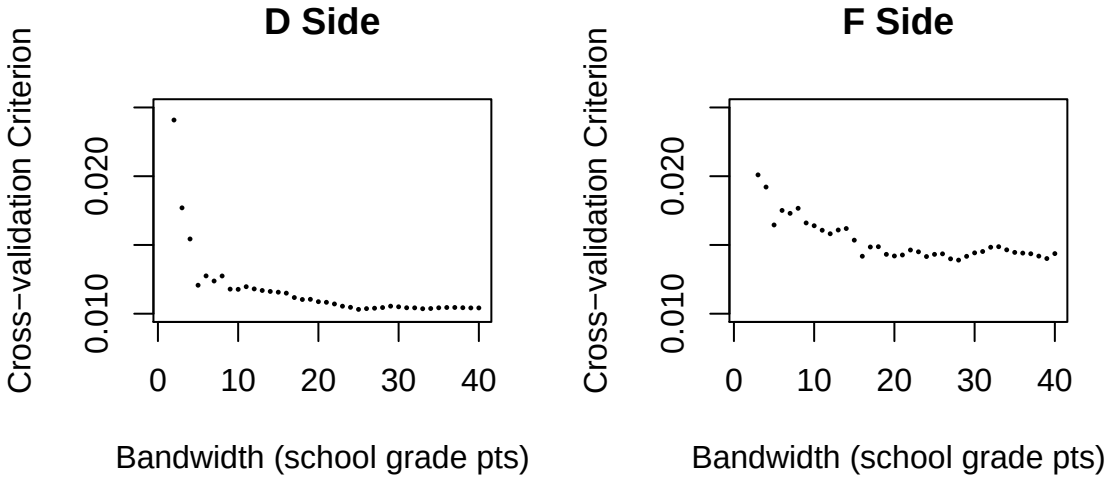


Table 9: Effect of 2002 School F Grade on Student Exit

Bandwidth	Point Estimate	Std Err	First Difference	Std Err
A-B				
5	0.885	9.316	0.025	0.618
10	1.695	4.984	0.148	0.506
20	0.071	0.185	0.009	0.173
30	0.079	0.145	0.008	0.015
B-C				
5	0.074	0.320	0.012	0.039
10	0.143	0.146	0.016	0.017
20	0.195	0.112	0.021	0.013
30	0.149	0.095	0.016	0.010
C-D				
5	0.139	0.215	0.021	0.030
10	0.085	0.187	0.012	0.024
20	-0.175	0.145	-0.021	0.018
26	-0.064	0.121	-0.008	0.015
30	-0.054	0.115	-0.006	0.014
D-F				
5	0.322	0.392	0.060	0.072
10	-0.241	0.330	-0.032	0.048
20	-0.130	0.215	-0.018	0.032
28	0.134	0.176	0.024	0.031
30	0.081	0.174	0.015	0.030

Notes: Estimates are from separate logit regressions. Standard errors are clustered by school. First differences reflect the increase in probability of exit of a B school grade relative to an A school grade. First differences are estimated through simulation while setting school grade points at the threshold value, student test scores at their mean value, and student demographics at their mode values.

Figure 9: A-B Threshold Cross-Validation Criterion for Entrance

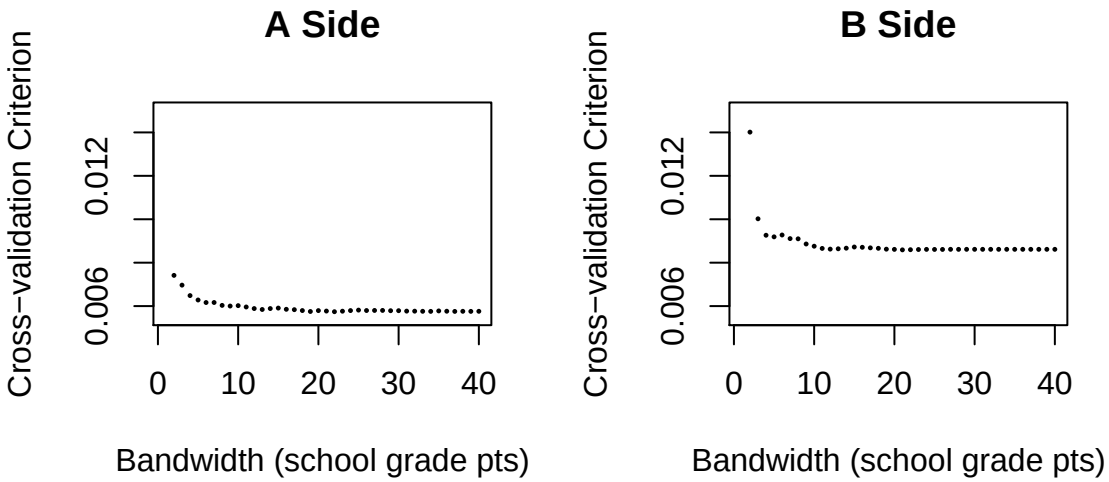


Figure 10: B-C Threshold Cross-Validation Criterion for Entrance

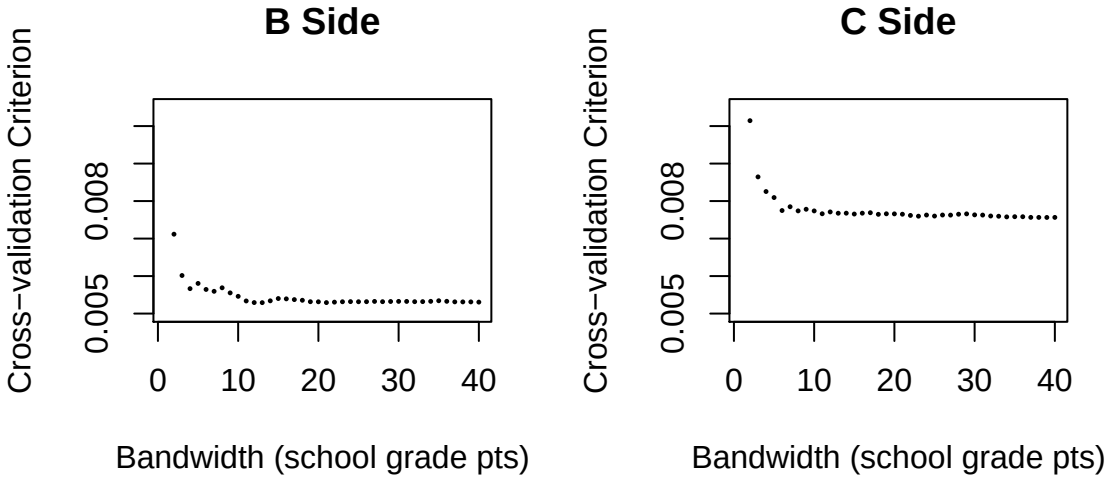


Figure 11: C-D Threshold Cross-Validation Criterion for Entrance

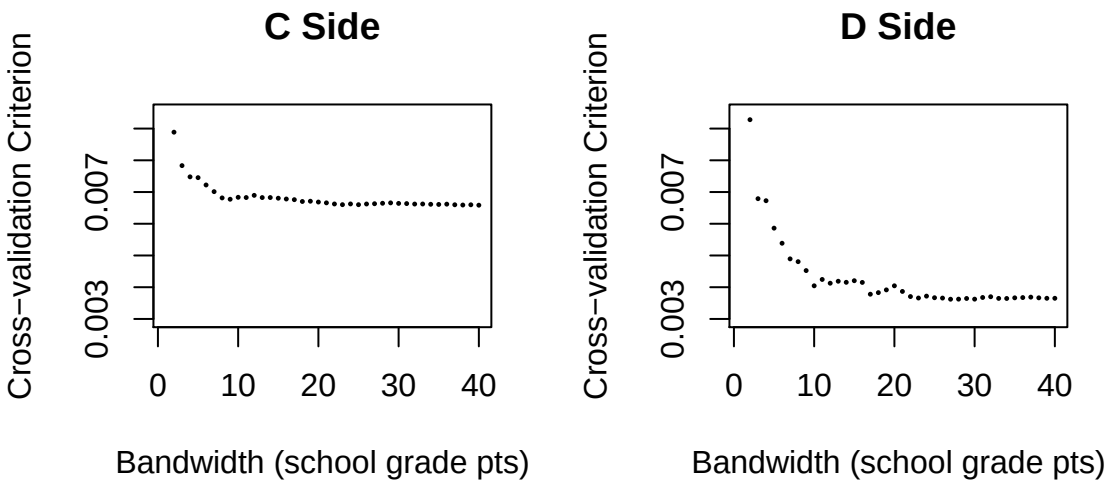


Figure 12: D-F Threshold Cross-Validation Criterion for Entrance