CENTRE for ECONOMIC PERFORMANCE

CEP Discussion Paper No 907

January 2009

To Leave or Not to Leave? A Regression Discontinuity Analysis of the Impact of Failing the High School Exit Exam

Dongshu Ou

THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

Abstract
The high school exit exam (HSEE) is rapidly becoming a standardized assessment procedure for educational accountability in the United States. I use a unique state-specific dataset to identify the effect of failing the HSEE on the likelihood that a student drops out early based on a Regression Discontinuity design. It shows that students who barely fail the exam are more likely to exit than those who barely pass despite being offered retest opportunities. The discontinuity amounts to a large proportion of the dropout probability of barely-failers, particularly for minority and low-income students, suggesting that the potential benefit of raising educational standards might come at the cost of increasing inequalities in the educational system.

Keywords: high school exit exam, student dropout, regression discontinuity
JEL code: I21, I28, J24

This paper was produced as part of the Centre's Education and Skills Programme. The Centre for Economic Performance is financed by the Economic and Social Research Council.

Acknowledgements
I thank New Jersey Department of Education for providing the data used in this paper. I am grateful to Dana Berry, Timothy Peters, Rob Akins and Don White for their help in answering my questions on the data as well as New Jersey's state educational assessment. I thank Henry Levin, Francisco Rivera-batiz, Randall Reback, Miguel Urquiola for valuable discussions and comments. For their helpful suggestions, I also thank Linda Bell, Alex Eble, Yinghua He, Dong Wook Jeong, Ayako Kondo, Stephan Litschig, Sandra McNally, Steve Pischke, Emma Tominey, Jessica Wolff and participants at Eastern Economic Association Meetings Annual Conference 2008, American Education Finance Association Annual Conference 2008, Campaign for Equity Student Research Conference, and Institut d'Economia de Barcelona's summer school 2008, Society of Economics and Education Seminar. The research is supported by the Campaign for Educational Equity Student Research Grant, the Spencer Research Training Grant and the Policy and Research Fellowship at Teachers College Columbia University. Earlier drafts of this paper are titled "The Impact of High School Exit Exam on Student Dropouts: A Regression Discontinuity Analysis." All remaining errors are mine.

Dongshu Ou is a Research Economist at the Centre for Economic Performance, London School of Economics.

Published by
Centre for Economic Performance
London School of Economics and Political Science
Houghton Street
London WC2A 2AE

ISBN 978-0-85328-340-9

1. Introduction

Many states in the U.S. require high school students to pass an exit exam as a graduation requirement. The exit exam is rapidly becoming a standardized assessment procedure for educational accountability. In 2006, high school students in 22 states were required to pass an exit exam to obtain their high school diplomas. Most states are phasing in exit exams and some have implemented more rigorous exams than the more traditional Minimum Competency Exams (MCE), adopting Standards-Based Exams (SBE) and End-of-Course exams (EOC)[1]. These High School Exit Exams (HSEE) are more prevalent in states with higher percentages of economically disadvantaged and minority students. Not surprisingly, the graduation rates of black and Hispanic students and students from low-income families are much lower than those of white students (Center on Education Policy[2], 2006).

HSEEs were introduced to verify that graduating students in high school had mastered the core curriculum (CEP, 2004). However, it has been suggested that such high stakes exams have adverse impacts that prompt minority and low-achieving students to leave high school early (CEP, 2006; Garcia and Gopal, 2003; Jacob, 2001; Warren, Jenkins, and Kulick, 2006). Furthermore, despite the vast and rapid adoption of the policy, there is very little causal research on the benefits of the HSEE including whether the exam effectively raises students' academic skills. On the other hand, most of the existing literature on HSEE show mixed evidence on the association between state HSEE policy and state level dropout and graduation rates[3]. Martorell (2004) first explored the causal relationship between failing the HSEE and various student level academic outcomes. He concluded that the HSEE in Texas in the 1990s does not discourage test failers to drop out *early*, but failing the exam[4] reduces post-secondary attainments.

This paper presents new empirical evidence[5] on whether failing the high school exit exam increases the chance of exiting from high school *prior to high school completion*. More importantly, I discuss the potentially different impacts of failing the HSEE on students with limited English proficiency, racial minorities and low-income students. This paper also intends to quantify the magnitude to which different testing subjects might affect student dropouts that previous studies have not addressed. I take advantage of a new longitudinal data

[1] These are three types of HSEE. MCE focus on basic skills below the high school level and SBE and EOC are aligned with high-school-level standards. Details are in the next section.
[2] CEP hereafter.
[3] Some papers use student level data such as Jacob (2001) and Griffin and Heidorn (1996).
[4] Martorell (2004) only finds statistically significant results in the last retest.
[5] There is concurrent and independent work on the same topic which I have recently become aware of. Please see literature review for a detailed description.

set from the State of New Jersey that captures the most recent changes (i.e. higher standards) in the exit exam. In particular, following Martorell (2004), I exploit the discontinuity in the likelihood of exiting around the cutoff score of HSEE, and compare the exit probability of the students who barely pass or barely fail the test. Barely-failers will provide the counterfactual outcome for barely-passers since the treatment status will be "as good as randomly assigned" in a neighborhood of the treatment threshold (van der Klaauw, 2002).

I demonstrate that students who barely failed the initial HSEE are more likely to exit high school early than students who barely passed. The difference in dropout probability among those who fail narrowly and those who pass narrowly is larger for racial minority students, economically disadvantaged students and for math tests relative to English tests. The estimates for the first test amount to a large proportion of the raw probabilities of drop out after initial failure of the exam, especially for math. The results are robust when using different functional forms to predict the discontinuity as well as to test the discontinuity in a small neighborhood around the cut-off score.

My analyses investigate the propensity to drop out for students around the pass/fail cut-off in HSEE. The difference in dropout propensity between the two groups can possibly be due to the psychological effect of failing (a "discouragement effect"), misunderstanding or not being informed of the retest opportunities, or withdrawal because of the high perceived cost of studying for the retests. While there is no causal evidence on the potential benefits of raising educational standards by HSEE, the identified dropouts found in this study suggest that high stakes testing has a potential tradeoff. Schools and policy makers should consider providing counseling services[6] for students who fail initially and better inform these students, especially minority students, about the availability of retest opportunities as well as reduce the stigma of failing HSEE (Cornell et al., 2006)[7].

This paper is organized as follows. Section 2 provides a background to the High School Exit Exam in the U.S. followed by a brief review of related literature. Section 3 introduces the exit exam in New Jersey. Section 4 describes the data and Section 5 discusses the empirical strategy used and its validity. Section 6 reports the empirical results, and Section 7 concludes the paper.

2. Overview of the High School Exit Exam in the U. S.

[6] McGray, Douglas. 2006. "Counseling Kids to Graduation and Beyond". Los Angeles Times. September 6. http://www.newamerica.net/publications/articles/2006/counseling_kids_to_graduation_and_beyond_4011

[7] Wilson, Latricia. 2008. "Exit Strategies: Confronting Faulty Grad Tests." WireTap Magazine. July 3. http://www.wiretapmag.org/education/43620/

"State high school exit exams are state mandated-tests that high school students must pass to receive a high school diploma"[8]. An exit exam is designed to test all subject matter learned during high school in a comprehensively fashion. The introduction of the HSEE as a graduation requirement can be traced to the 1980s[9] when the report *A Nation at Risk* (1983)[10] called for higher standards and expectations including the use of standardized tests to improve the academic underachievement of U.S. students compared to other advanced countries (Dee, 2002; Dee and Jacob, forthcoming; Harris and Herrington, 2006; Warren et al. 2006).

Existing HSEEs can be categorized as the following three types: (1) Minimum Competency Exams (MCE); (2) Standards-Based Exams (SBE); and (3) End of Course (EOC) Exams. The MCE is a test that focuses on basic skills below the high school level, and is usually administered in 9th or 10th grades. The SBE is aligned with state standards and is targeted at the high school level. It varies from state to state in terms of which grade is tested initially. The subjects tested in MCEs and SBEs are usually reading, writing and math. The EOCs are tied to a specific course at the high school level. They measure knowledge of each subject separately and are taken immediately after students complete the coursework in a given subject.

States with current exit exams allow students who do not pass the exam on the first try to retake it before the end of 12th grade and even after completing 12th grade. Strictly speaking, students are not able to graduate from high school if they fulfill other state or local requirements (such as coursework) but score lower than the "cut-off" score, i.e. the required proficiency level, in the HSEE. However, some states have alternative paths for students who do not pass the exams to graduate from high school (CEP, 2006)[11].

The federal *No Child Left Behind Act* (NCLB) of 2001 set out to reform public education by introducing accountability measures, which require that each state administer annual standards-based assessments to students in grades 3 through 8, and at least once in high school. Most state exit exams are thus used to meet the NCLB high school assessment

[8] Center on Education Policy. 2006. http://www.cep-dc.org (assessed December 19, 2008).
[9] High stakes testing was first integrated into student accountability reform in the 1970s and became more common in 1980s (Warren et al. 2006). For a detailed discussion on the educational reform movement, see Dee (2002), Dee and Jacob (2006), Harris and Herrington (2006), Warren et al. (2006).
[10] National Commission on Excellence in Education (1983). "A Nation at Risk".
[11] In New York, for example, students can use certain Advanced Placements (AP), International Baccalaureate and SATII tests as substitutes for the State Regents exams (NJDOE, 2006). In New Jersey, there is a Special Review Assessment which allows students to graduate in replacement of the exit exam. See a detailed discussion on New Jersey in next section.

requirements[12] (CEP, 2006, 2007). As NCLB reinforces and expands the requirements for high school students to demonstrate a level of competency in order to graduate from high school, the HSEEs also developed rapidly in recent years in terms of educational standards: either moving from MCEs to SBEs or EOCs, or increasing the test difficulties or requirements[13]. In addition, HSEEs expanded quickly to different regions. By 1992, MCEs had been adopted by every southern state except Arkansas and Oklahoma (Bishop, 2005) and fifteen states required the class of 1992 to pass the high stakes testing requirement (Jacob, 2001). Currently there are sixteen states using SBEs and four states using EOCs for their high school graduation requirement. Only three states retain the MCE(CEP, 2007). Table A-I presents some characteristics of the most current state HSEEs including type, subjects tested, first tested grade and the first graduating class affected by HSEE.

2.1 Previous Literature Related to the HSEE and Student Dropouts

As more and more states are adopting the HSEE as their graduation requirement, and education stakeholders express their concern about the policy consequences, there is increasing interest in studying the impact of the HSEE at both the national and state levels. The Center on Education Policy (CEP) has published six annual reports on the HSEE since 2002. The reports summarize a broad range of literature, both qualitative and quantitative, on the impact of the HSEE on various student outcomes (CEP, 2004, 2006). A large literature has been focused on dropout rates, not only because the dropout rates in HSEE states tend to be high[14] but also because the cost of dropping out is substantial both to the student and to society. For example, one recent study by Belfield and Levin (2007) has shown that the estimated social gains for an additional graduate could be up to $392,000 in present value of a person who aged twenty in 2007. For the purpose of this paper, I will briefly describe below some recent quantitative papers related to the HSEE and student dropout rates.

Opponents of exit exams (Griffin and Heidorn, 1996; Jacob, 2001; Warren et al. 2006) often claim that exit exams encourage more students to drop out[15]. For example, Warren and his colleagues (2006) used a combined data set of Current Population Survey (CPS) and Common Core Data (CCD) to measure the impact of high school exit exam policy on number

[12] For some states the passing score to receive a high school diploma is lower than the proficient level demanded by NCLB. See a detailed survey by CEP (CEP, 2006).
[13] However, such exams measure learning in grades earlier than grade 12 and thus might not measures of what is supposed to be learned overall in high school.
[14] Though New Jersey has a relatively low rate compared to other HSEE states.
[15] Griffin and Heidorn (1996) find an increased likelihood of dropping out for higher GPA students associated with failure on Florida's competency test, but not for minority or low achieving students. Jacob (2001) finds that low ability students are more likely dropout in states with exit exam policy than those without using National Education Longitudinal Survey data.

of high school completers. Comparing results using three different definitions of high school graduates, the authors provide evidence that states with more difficult HSEEs tend to have lower high school completion rates[16]. However, other studies show no evidence of a relationship between exit exams and dropping out (Greene and Winters, 2004; Muller, 1998; Muller and Schiller, 2000). For instance, Greene and Winters (2004) use a fixed effect model and two calculations of graduation rates for the class of 1991 to the class of 2001 to evaluate the impact of state HSEEs on graduation rates. The coefficients on HSEE policy dummies are small and insignificant, from which the authors concluded that adopting a HSEE has no effect on high school graduation rates. They claim that it is possible that the increased dropout rate of those who fail exit exam offset the increased graduation rate of those who are motivated by the exam.

A more recent study done by Dee and Jacob (forthcoming) reported that Minnesota's exit exam increases the dropout rate in urban and high-poverty school districts as well as in those with a relatively large concentration of minority students. Dee and Jacob (Forthcoming) also conclude that students in states with exit exams are more likely to drop out of high school than those students of similar backgrounds in states with no exams. The effects are strongest among African-American males.

These studies reach different conclusions mainly due to data differences and different strategies for identifying effects. For example, the data sets vary from the National Educational Longitudinal Study of 1988 (NELS: 88) (Jacob, 2001; Mueller, 1998; Mueller and Schiller, 2000) to the Census (Warren and Edwards, 2005; Warren et al., 2006) and the Common Core Data (Dee and Jacob, Forthcoming). Some of these studies encounter data limitations and the calculation of student outcomes varies considerably[17] in these data sets. The Census, for example, does not distinguish GED and regular diploma recipients. The NELS data set, on the other hand, is based on MCEs in the 1980s, which differ significantly from the current HSEE. It also limits the analysis to a single cohort. In addition, though the national survey and Census both allow cross-state analysis, they only give an average estimate of the policy impact. Because the high school exit exam is set at a state level, the nationally-based analysis does not capture the substantial variation across the states. The methodologies used in previous studies, such as cross sectional analysis or the fixed effect model, are also limited in establishing causality (Loeb and Strunk, 2003). It is not clear, as

[16] They also find it associated with a higher percentage of taking the General Education Development tests.
[17] See, for example, Lawrence Mishel (2006) "The Exaggerated Dropout Crisis," *Education Week*, March 8, for a discussion on calculation of graduation rate from CPS, NELS and CCD.

advocates of the HSEE argue, whether "many of the students who don't pass exit exams would have failed to graduate anyway" (Greene and Winters, 2004). Without a prospectively randomized trial or a convincing identification strategy for this retrospective data, we could not rule out the fact that the detected relationship between HSEE and student outcomes may be due to some unobservable factors which would have affected graduation rates irrespective of HSEE requirements.

Using a regression-discontinuity method on a longitudinal dataset, Martorell (2004) studied the impact of the Texas exit exam during the 1990s (Texas Assessment of Academic Skills, TAAS) on several educational outcomes including dropout and graduation rates. Comparing the students who barely pass or barely fail the TAAS, he concluded that the test does not "discourage" students to drop out in early grades, but failing the exam reduces the amount of post-secondary attainment. However, the insignificant and minute impact on student dropouts could be due to the multiple retesting opportunities for the TAAS[18] and that the TAAS measures more basic skills than the newer exam, the Texas Assessment of Knowledge and Skills (TAKS).

The "discouragement effect" resulting from failing the HSEE is also discussed by Cornell et al. (2006). Their qualitative study documents the potential psychological effect driven by failing the HSEE[19], which might discourage students from completing high school, leading to early exit.

While multiple subjects are tested in the existing HSEEs, few studies have been done to quantify the magnitude to which failing different subjects might affect student dropouts, especially for minority and low-income students. This paper intends to fill that gap in the current literature by examining a specific state within a given period of time. In the process of writing this manuscript, I became aware of two concurrent and independent studies that also address this important question on which we have relatively little evidence. Papay, Murnane, and Willett (2008) focus on the on-time graduation rate for barely-failers relative to barely-passers of Massachusetts's HSEE (Massachusetts Comprehensive Assessment System, MCAS) in urban schools in 2006. They find some negative effect of failing the math test of MCAS for the low-income urban students, however, no significant effect for failing the English test. A recent manuscript by Reardon and his colleagues (2008) looks at the effect of failing California's HSEE (CAHSEE) on student's subsequent academic achievement and

[18] This is also mentioned in Greene and Winters (2004) and CEP (2004) on the reasons that researchers do not find significant results on increasing dropout rate associated with the implementation of state high stakes testing.
[19] The paper studies the students who were wrongly informed as failing the exam.

persistence to 12th grade as well as graduation rate. They find little evidence of CAHSEE failure on all outcomes measured in four large school districts in California.

Similar to these studies, which also use regression discontinuity design, this paper does not attempt to address the question of whether the state HSEE policy affects student academic outcomes differently. Rather, it provides evidence as to the impact of failing today's more rigorous exams on students around the pass/fail cut-off. Moreover, the large dataset allows me to compare results in four cohorts and different subgroups. The analysis also sheds light on how students react differently to their performance on initial test and retest.

3. The HSEE in New Jersey

New Jersey was one of the first states to adopt a statewide assessment test as a requirement for its high school diploma. All students in public schools in New Jersey must pass the HSEE as a graduation requirement. Early in 1981-1982, 9th grade students needed to pass the Minimum Basic Skills Test to get a high school diploma. The test has been upgraded to a more challenging assessment, i.e. the Grade 9 High School Proficiency Test (HSPT9) in 1983. Student who were first time 9th graders on and after the fall semester in 1991 were required to pass the High School Proficiency Test (HSPT 11) to graduate. The state began to administer its new standard-based exam, the High School Proficiency Assessment (HSPA), in March 2002 for the first time to 11th graders, as a replacement for its previous version of the HSEE, i.e. HSPT11. With the advent of the HSPA, the HSPT 11 was administered to adult high school and returning students only. The final administration of the HSPT 11 was scheduled for spring 2004. The class of 2003 was the first cohort to graduate under this new HSEE requirement.

The HSPA is intended to gauge students' knowledge and skills in the New Jersey Core Curriculum Content Standard and is mandatory for New Jersey high school juniors in public schools[20]. The test contains two sections[21]: Math and Language Art Literacy (LAL). Both sections contain multiple-choice questions as well as open-ended questions. The LAL section also includes essays. The multiple-choice questions are machine-scored by a company hired by the New Jersey Department of Education, and the open-ended questions and essays are scored by two trained independent raters whose scores are combined to calculate the total score. Each section of the test is scored separately. The scores on each section range from 100 to 300. The state defines a score under 200 as "partially proficient", 200 to 249 as

[20] http://www.state.nj.us/education/assessment/hs/.
[21] Science is included from Spring 2007 and "Social Studies" is tentatively scheduled for later testing under the HSPA umbrella.

"proficient" and 250 and above as "advanced proficient". A procedure called statistical equating is used to make sure that all future HSPA tests are at the same level of difficulty as the March 2002 test. In order to pass the HSPA, a student must obtain a score of at least 200 on all sections[22].

Eleventh graders are first tested on both math and LAL during the spring semester[23] and, subsequently, during the fall semester and spring semester in twelfth grade if they fail either section. Students only need to retake the test in the subject that they fail. Students have three opportunities to pass the exam before the completion of 12th grade[24]. "Through a statistical equating procedure, the HSPA scores [are] comparable from administration to administration (NJDOE, 2007)."

Students identified as having limited English proficiency (LEP) must meet the same graduation requirements as native English speaking students. For students with an educational disability, an Individual Education Program (IEP) defines their graduation requirements. If a student's IEP includes an exemption from passing any subject of the HSPA, then the student will need to take the subject at least once but does not need to have a passing score to graduate from high school.

4. Data

The data used in the following analysis are obtained from the New Jersey Department of Education (NJDOE), which included records of test scores on the HSPA of all students who enrolled in a high school in New Jersey from 2002 to 2006. In addition to Math and Language Arts Literacy (LAL) test scores and school enrollment information, the data set contains information on the student's school, grade, gender, age, race/ethnicity, whether s/he is economically disadvantaged, special education status, limited English proficiency, an indicator for less than 1 year spent in the school of testing, IEP exemptions and an indicator of socioeconomic status of citizens in each district.

My analysis is restricted to public school general education students who took HSPA initially in their 11th grade. I trace these students to the end of 12th grade and construct a longitudinal data set for four cohorts. I have excluded students with missing information,

---

[22] Students, who have met all the graduation requirements except passing the desired level (cut-off point) of HSPA, can undergo a Special Review Assessment (SRA). The SRA is aligned to the HSPA test specifications and is usually taken after school district personnel receive the specific information about proficiency results for each student.

[23] The testing lasts for two days and may be taken on alternate days if a conflict exists (such as religious obligation).

[24] For students who exited from high school (for any reasons including finishing 12th grade but did not receive the diploma), they can still take the exam as a "returning student" whenever the test is administered. Another option for students is to attend adult schools and retake the test.

migrant students and students who stayed in school for less than one year when the test was administered. The final sample I use in this paper consists of 299,948 observations for all cohorts[25]. A detailed discussion on how I construct the data and select the sample is provided in Appendix I.

Table 1 reports the descriptive statistics by year and for different subgroups. In general, there are about 51 percent female students and 66 percent white students. Black, Hispanic and economically disadvantaged students each comprise about 13 percent of the sample[26].

The test scores are adjusted by subtracting the cut-off score 200 from each score, so that greater than or equal to zero is passing and less than zero is failing. The initial passing rate of math test is lower than that for the LAL test. Female and white students perform better and have a higher initial passing rate than black, Hispanic and economically disadvantaged students. For instance, black students on average receive a math score under the cutoff (-9.5 with a standard deviation of 47.3), while white students score 28.15 point above the cutoff (with a standard deviation of 37.38). Not surprisingly, the average exit probability after the initial test is higher for black, Hispanic and economically disadvantaged students[27] than that for white students. Compared to other groups, the LEP students has an extremely low passing rate (less than 30% in either math or LAL) and a high exit rate after failing one or both subjects of HSPA.

5. Empirical Strategy

5.1 Regression Discontinuity Design and the Estimation of Failing HSEE on Dropping Out

I examine two situations in this paper. The first is how likely it is that students will exit after failing the initial attempt. The other is the likelihood of dropping out after the first retest. Denote $Exit_i^0 = 1$ if a student i exits after the semester of initial test while $Exit_i^0 = 0$ indicates

[25] Special education students with IEP exemption in any of the three test administrations are not included. The justification for this exclusion is that special education students who get IEP exemption can graduate from high school without passing the exams as long as other graduation requirements are fulfilled. There are also cases where some special education students have been exempted from taking the test because of a personal situation. Their HSPA test scores do not indicate a "pass" or "fail" status. This would not give a perfect assignment to the cut-off and violate the sharp RD assumption. I also excluded the special education students in the retest sample. See next section and Appendix for further discussion.

[26] New Jersey has the lowest rates of students eligible for free or reduced-cost meals (27%) according to NCES data (Mackey, 2006). My sample has a lower percentage of economically disadvantaged students because of the analysis only includes students who took HSPA the first time in 11th grade. My sample also has a higher percentage of whites, and lower proportion of Hispanic and Black students, whereas the State has 58% White, 18% Black and 17% Hispanic in 2003-04 (Mackey, 2006).

[27] Though the exit rate after first failure (second last row) is high for all groups.

the student stays on to the following semester of the initial test. The test score of subject *s* (s=math/LAL) of the initial test is denoted as $SCR^{0}_{i,s}$ where passing score is adjusted to 0 here. For student i who doesn't pass the exam at the initial attempt (i.e. $SCR^{0}_{i,s} <0$), he might choose to leave high school ($Exit_i^0 =1$), or stay on ($Exit_i^0 =0$) and take a retest on the failing subject(s) in the fall of 12th grade[28] and receive a score SCR^{1}_{s}. If the student fails again ($SCR^{1}_{i,lal} <0$ and/or $SCR^{1}_{i,math} <0$), he might choose to exit ($Exit_i^1 =1$) or stay on ($Exit_i^1 =0$) to take the second retest (i.e. last chance) in Spring of 12th grade. Retest on failing subject(s) is required by NJDOE as long as the student is enrolled.

Singling out the true effect of failing HSEE on a student's dropout behavior is difficult as the observational data may not provide sufficient information to capture the variations of reason for dropping out. For example, students who are less likely to fulfill other academic requirements (e.g. sufficient credits) tend to drop out regardless of passing the HSEE (Greene and Winters, 2004). The regression discontinuity approach helps to overcome the identification issues.

I employ a sharp regression discontinuity (RD) design (Trochim, 1984) to identify the effects of failing HSPA on exiting high school. There is a growing literature on educational and social program evaluations using the RD approach (Angrist and Lavy, 1999; Chay, McEwan and Urquiola, 2005; Jacob and Lefgren, 2004, 2007; Kane, 2003; van der Klaauw, 2002; Lemieux and Milligan, 2008; Martorell, 2004; McEwan and Shapiro, 2008). The RD approach allows me to compare the outcomes for students whose test score is "just below" and "just above" the cut-off score, since, on average, they will have similar characteristics except for the treatment.

Given that the individuals just above and below the cutoff are very similar in that they "have similar average outcomes in the absence of the [treatment] program as well as similar average outcomes when receiving treatment" (van der Klaauw, 2002), we may assume that students closed to threshold of the pass/fail cut-off score (-ε, ε) have the same dropout rate in the absence of "failing"status:

$$P(Exit_i^t = 1 \mid SCR_{i,s}^t) = \alpha \qquad \text{if } SCR^{t}_{i,s} \geq 0$$

$$P(Exit_i^t = 1 \mid SCR_{i,s}^t) = \alpha + \tau_{it} \qquad \text{if } SCR^{t}_{i,s} <0$$

The discontinuity for exiting hence follows:

$$D_{exit} = P(Exit_i^t = 1 \mid SCR_{i,s}^t = \varepsilon) - P(Exit_i^t = 1 \mid SCR_{i,s}^t = -\varepsilon) = \alpha - (\alpha + \tau_{it}) = -\tau_{it}$$

[28] I do not consider students who are retained in 11th grade in this study.

with t= 0 for initial test and t=1 for first retest. The difference between the exit probabilities for barely-passers and barely-failers, denoted as τ_{it} **,** is the "random shock of failing the test[29]" that this paper attempts to investigate. Note that the RD provides a local estimate for the subgroup of individuals around the cut-off point. This indicates that any effects estimated in the model only apply to the students who are around the threshold.

My main estimation equation is given by:

$$Exit_i^t = \gamma_1 * 1(SCR_{i,s}^t \geq 0) + f(SCR_{i,s}^t) + \gamma_2 * 1(SCR_{i,s}^t \geq 0) * f(SCR_{i,s}^t) + \eta_i^t \qquad (1)$$

where $f(SCR_{i,s}^t)$ is an unknown smooth function of test score. γ_1 is the coefficient of interest. The estimated discontinuity will be the marginal effect of passing HSPA obtained from the probit regressions. $1(\bullet)$ is an indicator function that is equal to one if the enclosed statement is true. η is the error term. To reduce the sampling variation, covariates will be added. This procedure will also test the credibility of the RD approach because adding the covariates should not change my RD estimates of the treatment effect if they are independent of the assignment variable (Imbens and Lemieux, 2008). Controls for t=0 include student and school baseline characteristics and dummies for different cohorts. For period t=1, I include an extra dummy variable for whether the student is retesting on only one subject along with the controls for t=0.

To illustrate the idea, I show the average exit rate after the initial and first retest in Figure A-I[30] . Both graphs show the relationship between the math/LAL scores (relative to cutoff) and the actual or estimated means of exiting high school using the pooled data[31]. The circles reflect the actual average probability of exiting at each score point. The solid line maps the predicted exit probability from a polynomial function in test score[32]. We can see the discontinuity of exit probability around the cutoff of the math and LAL test scores in both initial and the first retest. The probability of exiting from high school jumps up at the cut-off score for students who barely failed the test. The patterns are clearer in graphs of the math test. This gives an overview of my methodology and the possible estimates of the impact of failing HSEE based on the discontinuities around the cutoffs.

---

[29] Martorell (2004) calls it a "discouragement effect". It can be interpreted as the "net discouragement effect", as the discontinuity also exists when the students who barely pass boost their confidence and therefore increases the probability of staying in high school. In principal, we are interested in the potential outcome differences (τ_{it}) between the barely-passers and barely-failers.

[30] The graphs show only the observations within a small threshold around the cut-off.

[31] The pooled data includes four cohorts. I also did the plotting for each cohort separately and the patterns indicate a valid use of sharp discontinuity design (Trochim, 1984; van der Klaauw, 2002).

[32] The Probit estimation I used here takes a dummy for passing the (Math/LAL) test, a fourth order polynomial in the score (a cubic form in the retest score), and an interaction between the passing dummy and a linear term in the test score.

5. 2. Validity Test and Robustness

It is important to distinguish any effects that are caused by observed variables other than the treatment itself. The underlying assumption is that the only source of discontinuity in the probability of dropping out at the cut-off score is failing the particular exam. Graphically, there should be no jumps in these other observed variables at the cut-off based on the test score function (Imbens and Lemieux, 2008). Table A- II presents the RD estimates on a set of pre-determined characteristics of the students around the cut-off, including indicators for race/ethnicity, gender, limited English proficiency, special education and economically disadvantaged status. Most results do not show any statistically significant differences between the barely-passers and barely-failers of the HSPA. A few do show that there is a discontinuity around the boundaries; however, plotting[33] reveals that they do not pose a threat to identification.

Furthermore, the correct specification of the smooth function $f(SCR_{i,s}^{t})$ is the key to identify the true treatment effect in the RD method (Angrist and Lavy, 1999; McCrary and Royer, 2005; Trochim, 1984). To check the robustness, I compare results from various specifications of the test score $f(SCR_{i,s}^{t})$ and I also include the interaction terms to allow any slope changes. The findings are not sensitive to these functional form differences.

Besides testing alternate specifications for $f(SCR_{i,s}^{t})$, robustness checks include data restrictions to those close to the cut-off and changes in the set of control variables in the Probit estimation (Trochim, 1984). Given the standard deviations, and to ensure a comparable number of students on either side of the cutoff[34], I choose different bandwidths for estimates on the math and LAL tests separately. These different ranges will allow us to compare the RD estimates and sufficiently test their robustness. The results are similar when I control for observable individual covariates such as race/ethnicity, gender, cohort dummies, etc.

To examine the heterogeneity of the impact across the student population, I will run the RD estimate for different subpopulations (e.g. failing one or both subjects, gender, race/ethnicity, cohort/year) and test statistical significance across groups and over time.

6. Results

[33] The only two variables that seem to discontinue around the cut-off are the dummy variable for gender and for the middle SES group (graphs not shown). However, my inclusion of these variables does not significantly change the treatment effect estimates in the discontinuity sample, suggesting that neither of these variables are strongly correlated with the treatment status.

[34] Histograms of the test scores do not show that any sharp increase of numbers of students just below or above the cut-offs.

All tables report the parametric estimates of the discontinuity in probability of exiting after the initial test or retest. The estimates are reported separately by specifications based on the Math or the Language Arts Literacy score. All the standard errors are clustered at the score level (Card and Lee, 2008). Accordingly, unless otherwise stated, the numbers used in the rest of the paper will be the estimates based on cubic form of $f(SCR_{i,s}^{0})$ or quadratic form of $f(SCR_{i,s}^{1})$ with the slope of these functions to vary on each side of the cutoff. For simplicity, only the estimates with covariates are discussed below.

In general, my RD estimates are comparable to and consistent with the discontinuities illustrated in the graphs (Figure A-I). The apparent discontinuity of exiting likelihood at the cut-off score suggests that students who barely fail the test are much more likely to drop out than students who barely pass. Overall, the discontinuities based on the Language Arts Literacy (LAL) scores are observed with smaller magnitude than those based on the Math scores. My results are robust to various functional forms[35] of the test score, to different bandwidths and to the inclusion of a set of baseline characteristics of the students.

6.1. Dropout and Failing the Initial Test

The estimates in Table 2 provide new evidence regarding the impact of failing HSEE on student early dropout behavior. The statistically significant RD estimates confirm the link between failing the initial test and the increased probability of dropping out. Results are robust either using different test score functional form or adding covariates in the estimation model.

The discontinuities based on the Math score are larger than those based on the Language Arts Literacy (LAL) score. For example, the RD estimates on the math score are twice as high as the RD estimates on the LAL score. The discontinuity in the probability of dropping out after the first test is -.011 with standard error of .001 for students who are around the cut-off on the math test. It is -.005 with standard error of .002 for students who barely failed or passed the LAL test. This pattern holds when looking at the observations in a smaller neighborhood closer to the cut-off score (Column 5, 6, 7 in Table 2). For example, within the bandwidth of 50 points above and below the cut-off, the estimated discontinuity based on the math score is -.009 with a standard error of .001, whereas the estimated discontinuity is -.003 with a standard error of .001 based on the LAL score.

[35] I have tested various specifications including linear, quadratic, cubic and fourth polynomial functional form of the math or LAL score, with and without controls of student baseline characteristics. For the ease of reporting, I do not include all the specifications in the tables.

The difference in RD estimates for math and LAL test implies that barely-failers on the math test have a higher propensity to drop out than barely-failers on the LAL test. We notice that the average passing rate in math is lower than LAL on the HSPA exam (76% vs. 87% as in Table 1): which is plausible that the math standards in HSPA is set too high or the students are poorly prepared for the math test[36]. If this is true, students might become more discouraged when failing a more difficult test, or that their perceived cost of preparing for (and passing) the retest is higher when there is no adequate educational support in building up their math skills.

The cells ("Exit after failure") under each bandwidth also show the actual average exit rate for test failers within each bandwidth. For instance, the average exit rate for student score within (-30,0) in math is 4.8% with a standard deviation of .214. Hence, the -.012 estimated discontinuity in exit probability equals to a 33.3%[37] increase in the mean exit rate of barely-failers without the random shock. It counts for 5.6% of the standard deviation of the raw exit probability of the barely-failers. Similarly, the raw probability of dropping out after failure of LAL test within bandwidth of (-25,25) is 6.1% (with standard deviation .239). The estimated discontinuity is -.007, which raises 13.0% of the raw dropout probability of the barely-failers and amount to 2.9% of the standard deviation.

Table 2 also presents RD estimates individually for each cohort. Except results for Cohort 1[38], the statistically significant coefficients in three cohorts show consistent evidence to the estimates obtained from the pooled sample. The difference between the RD estimates based on math and LAL score is seen the largest in Cohort 4. The estimated probability of exiting for those barely failing in math is 2.2% (standard error .002)higher than those barely passing the exam in math, and .4% (standard error .002) higher for those barely failing LAL test compared to barely passing LAL test. It is worth noting that students in 2005 (i.e. Cohort 4) were given more time to take the math test while the number of testing items and difficulty remained the same. It is possible that the psychological impact of getting discouraged rises if students who believe they can achieve a higher score with the extended testing time, in fact, still fail.

---

[36] If failure of students who barely fail or pass the test is solely due to measurement error (for example, poor construction of the test), and that the error is bigger in LAL test than that in math test, then students will be more likely to stay on if they fail only the LAL section than if they fail only the math section. However, the standard errors in the test scores (Table 1) shows that the variances in the math test is actually bigger than the LAL test, which implies that the bigger discontinuity in math test score specification does not result from testing error.

[37] Similarly, the percentage of increased exit probability increased resulted from the shock in column 6 and 7 is 22.2% and 18.4%.

[38] This is the first graduating class that was required to pass HSPA and the estimates might be noisy. I will focus my discussion on cohort 2, cohort 3 and cohort 4.

6.2 Dropout and Failing the Retest

The discontinuities of exiting at the cut-off after the second test are shown in Table 3. The estimated magnitudes are similar to the results of the initial test. For instance, the discontinuity estimate of exiting among math retest-takers is -.013 with a standard error of .001 for all cohorts. Barely-failers in math retest are more likely to drop out than barely-failers in LAL retest when compared to their counter groups (i.e. barely-passers in math/LAL test). Some results based on the LAL retest scores in each cohort are not statistically significant, which may be due to the reduced sample size.

However, my estimates might underestimate the impact because the retesting sample is different than the first-test sample. It excludes students who dropped out after failing the first test. Students who stay on for the second test are assumed to have higher persistence or stronger motivation to pursue their academic training and thus are not as likely to dropout even if they fail the second time. Therefore, the similar students dropout probability around the cut-off between the initial test and retest found in my results, based on data taken from New Jersey, could possibly reflect evidence of “discouragement effect” found in Martorell’s (2004) paper. He observes no statistical significance of “discouragement effect” of failing HSEE except for the very last retest in 12th grade. He explained that the discouragement effect was possibly due to the fact that students were unable to graduate from high school if they failed the test on their last chance to take it. In addition, seeing from the positive average test score gains from initial test to retest for the barely-failers around the cut-off[39], it is plausible that the discouragement is big if barely-failers of the initial test put great efforts in preparing for the retest but still fail.

It is also noteworthy that the discontinuity amount to an even larger proportion of the actual exit probability of the barely-failers in the retest. For example, the average exit rate for math barely-failers is 2%, and the RD estimates indicate an increase of 1.5% of the exit rate compared to the barely-passers in math retest.

6.3. Heterogeneous Effects

Table 4 and Table 5 illustrate the RD estimates for the sub-samples on initial test and retest respectively. Apparently, barely-failers in racial minority and economically disadvantaged groups are more likely to drop out around the cut-off level, especially after the initial test. For Hispanic and economically disadvantaged groups, the RD estimate is -.013

---

[39] For example, for math retesters scored below cut-off for less than 30 points, the average adjusted score in the first test is -31.22 (with a standard error of 44.48), and their average adjusted score in retest is -17.34(8.24). For LAL retesters scored below cut-off for less than 30 points, the average adjusted score in the first test is -7.93 (with a standard error of 41.21), and their average adjusted score in retest is -14.57(8.68).

with a standard error of .002 on the math initial test. This is twice as bigger as the discontinuity found in the white students (the RD estimate is around -.005 with a standard error of .001).

Interestingly, the dropping out pattern for retest seems to be different. Female and white students who barely fail the retest are comparable to those racial minority and economically disadvantaged students. Given that the average exit rates for female and white students are lower than those for the other subgroups[40], the discontinuities found in these two groups could amount to a larger proportion of the actual dropout probabilities. However, one major concern is that racial minority and economically disadvantaged students drop out earlier – since they are more likely to exit at the failure of the first test. As previously noted, the sorted retest sample for these deprived students, due to earlier dropouts after the initial test, could underestimate the treatment effect.

Lastly, my results do not confirm a statistically significant impact for LEP students of barely failing the exam. However, the descriptive statistics in Table 1 Panel B still show some support to the claim that there is potential negative impact of HSEE policy on students whose first language is not English (Garcia and Gopal, 2003). From 2002 to 2006, the average passing rate among LEP students is 28% for math's initial test and 22% for English initial test. The average math test score is -20.06 (with s.e.=45.05) and average LAL score is -33.37 (with s.e. = 43.96). Further, 8% of students who failed at least one subject in HSPA dropped out, which is higher than the average. As Garcia and Gopal (2003) describe, the LEP students are very likely to exit because the exam makes them ineligible to participate in the rigorous curriculum.

6.4. Effects of failing at least one subject

Are barely-failers more likely to drop out if they fail the other subject too? Estimates in Table 4 are performed for a restricted sample who failed the subject test other than the one used in the test score specification. Most of the results, especially estimates for the initial test, are not statistically significant. The results on the retest for all cohorts are consistent with the estimated magnitudes in the unrestricted sample. This indicates that the barely-failers are as likely as the barely-passers in HSPA to stay or drop out from high school if they fail both subjects[41].

---

[40] The average exit rate for females, LEP students, whites, Hispanics, blacks, and low-income students are .022, .040,.020, .032,.027 and .031 respectively.

[41] Nevertheless, barely-failers on either math or LAL who failed the other subject and exited high school have a lower score on the other subject (Results not shown).

On the other hand, given the failure of one subject, barely failing another subject means the failers have to prepare for both subjects in the retest. Regardless of the existence of Special Review Assessment as an alternative to graduate in New Jersey, the bigger time constraint that barely-failers will be facing in the final semester of high school might simply push them to drop out.

7. Conclusions and Policy Implications

The increasing movement towards school accountability is associated with the increased use of standardized tests. The High School Exit Exam is controversial in its stance on whether or not tests such as these could stimulate students' motivation and enhance learning or prevent more high school students from graduating. There is also concern that the potential gain of high stakes testing comes at the cost of increasing inequality between groups. This paper examines whether such a test can potentially affect students' decision on completing high school by comparing students who are very close to the pass/fail cut-off and compare the impacts for different subgroups including racial minorities and economically disadvantaged students.

Using data from New Jersey High School Proficiency Assessment (HSPA) exam, the paper finds statistically significant evidence that students who barely fail the exam, especially in math subject, are more likely to drop out than students who barely pass it. The similar effect is found in both the initial and the retest. While the raw dropout rate of the barely-failers in the initial test is 5% to 6%, the RD estimates indicate a 1% increase of the exit probability for barely-failers compared to barely-passers. In the retest the RD estimates indicate a 1% to 1.5% increase of exit probability for barely-failers whose actual exit rate is 2% to 3%. Though there is no significant evidence that Limited English Proficiency (LEP) students, the regression discontinuity estimates in dropout propensity are larger and statistically significant for black and Hispanic students as well as economically disadvantaged students.

My results are different than the previous study on this topic by Martorell (2004), who uses Texas data, as he does not observe a significant discouragement effect, even for students who fail a number of times. There are several explanations that might illustrate my results. One trivial reason is that the HSEEs we examine are very different and set under a very different context. Not only are the designs of the tests (both their contents and the number of retakes) different, but the time periods we study differ as well. Using recent data from Massachusetts, Papay et al. (2008) also discover that failing the math test in HSEE reduces the graduation rate of economically disadvantaged students. Though another recent study by Reardon and his colleagues (2008) does not find any evidence on failing California's HSEE on students' graduation rate for low-income and Hispanic students, their results might be limited by the district level data, which could not capture students' mobility within the state[42].

The larger measured impact of failing the test on racial ethnicity groups and economically disadvantaged students found in this research suggests that states should consider targeting these students and providing additional academic support to reduce dropouts among their ranks. Students who barely fail might be more likely to drop out than those who barely pass if they try hard but still fail and get discouraged; or the perceived cost of preparing the retests is high when not much remedial resources are available; or schools fail to give sufficient information on their retake opportunities and alternative way of graduation. Therefore, it is imperative for schools to make sure students are aware of the alternative mechanism of getting a high school diploma (Such as New Jersey's SRA) in the event of test failure.

That said, if the students who barely failed the HSEE are no different than the students who barely pass, school counselors and teachers should encourage the students to retake the test, and reduce the stigma of retesting to offset any "discouragement" effects (Cornell et al. 2006)[43]. Furthermore, other guidance and remedial courses might reduce students' perceived cost of retaking the test and increase the probability of staying in high school if they fail (Reardon et al., 2008).

Though the discontinuity of exit probability is similar in both the first and second test, it is worth noting that the magnitudes estimated in the retest may reflect sample selection. Students who drop out before the first test administration are not included in the sample[44]. In addition, students who were enrolled in 11th grade are less likely to drop out given that they already have a higher academic persistence compared to those who drop out in earlier.

One caveat is that the RD analysis provides local estimates and does not evaluate the overall effects of the exit exam policy on students' dropout behavior. Nevertheless, this study expands the scope of research on high-stakes testing and its association with racial minority and low income students. The results have a broader implication for the U.S. and are suggestive of the impacts that are caused by failing the HSEE. New Jersey recently launched

---

[42] See discussion in Reardon et al. (2008).
[43] See footnote 6 and footnote 7.
[44] If students were pre-screened and selected out of the testing pool by the schools or teachers (Reback, 2008), they would not be in my sample either.

the "High School Graduation Campaign" as part of a national effort to reduce dropout rates[45]. It is crucial for schools, parents and policy makers to understand the dropout decision process behind the choice to drop out and make use of the exit exam test scores to inform teachers and schools in improving their curriculum and to offer improved educational support to prevent dropouts. It is important to note that despite the rapid and vast adoption of the HSEE across US States, there is no sophisticated benefit analysis or evaluation of the policy to provide a solid background for the reform.

Test-based accountability may be effective at raising student achievement by motivating students and teachers to work harder and forcing school administrators to implement more effective instruction (Bishop, 1998; Carnoy and Loeb, 2002 ; Fuller and Johnson, Jr. 2001; Hanushek and Raymond, 2004, 2005; Jacob, 2001). Besides the consequences of failing on dropping out (Martorell, 2004, Papay et al., 2008), there might be other gaming devices that will undermine the substantive standards towards which the tests are designed: for example, incentives for schools to cheat (Jacob and Levitt, 2003), teach to the test rather than building knowledge and skills (Hoffman, Assaf and Paris, 2001; Jones, et al., 1999; Lazear, 2004), induce the selection of students with the goal of artificially raising scores (Cullen and Reback, 2006; Figlio and Getzler, 2002), or raise the passing rate focus educational resources on students close to the cutoff (Reback, 2008). Without more data and further evidence, we cannot argue whether imposing a higher standard of testing increases or decreases students' academic learning, nor can we allege that the cost of causing the marginal students to drop out outweighs the benefit of increased student achievement of the overall population.

Future research can be done by linking the HSPA test scores with information on other test scores (e.g. the SAT score), Special Review Assessment, and post high school indicators. The SRA information can test the assumption that students in a school with more students graduated via SRA are less likely to drop out if they fail the HSPA. Furthermore, it is important to know whether students who drop out early will drop out permanently. The cost of permanent dropouts is clearly higher than that of temporary dropouts (Belfield and Levin, 2007).

[45] Hu, Winnie. 2008. "A Plan to Cut the High School Dropout Rate." *New York Times*. October 24. http://www.nytimes.com/2008/10/26/nyregion/new-jersey/26educnj.html?_r=1&ref=education&oref=slogin

TABLE 1: Descriptive Statistics

Panel A. Student Characteristics By Cohort

Variables	Cohort1	Cohort 2	Cohort 3	Cohort 4	All cohorts
Famale	.51	.51	.51	.51	.51
	[.50]	[.50]	[.50]	[.50]	[.50]
White	.66	.66	.66	.64	.66
	[.47]	[.47]	[.47]	[.48]	[.47]
Black	.13	.13	.13	.13	.13
	[.33]	[.33]	[.34]	[.33]	[.34]
Hispanic	.12	.12	.12	.13	.12
	[.33]	[.33]	[.33]	[.33]	[.33]
Economic Disadvantaged	.12	.12	.13	.13	.13
	[.33]	[.33]	[.34]	[.34]	[.34]
Limited English Proficiency	.03	.03	.03	.03	.03
	[.18]	[.18]	[.17]	[.16]	[.17]
Special Education Students	.07	.07	.06	.07	.07
	[.25]	[.25]	[.24]	[.25]	[.25]
Age	17.02	17.02	17.01	17.00	17.01
	[.56]	[.56]	[.55]	[.53]	[.56]
Average Math Score (adjusted)	16.10	16.50	21.30	25.45	19.98
	[43.83]	[42.08]	[41.32]	[42.80]	[42.69]
Average LAL Score (adjusted)	21.86	22.77	25.46	27.65	24.52
	[4.07]	[38.79]	[35.52]	[34.82]	[37.36]
Initial Passing Rate: Math	.73	.71	.76	.81	.76
	[.44]	[.45]	[.42]	[.39]	[.43]
Initial Passing Rate: LAL	.85	.85	.88	.89	.87
	[.35]	[.35]	[.32]	[.32]	[.34]
Exit (11th grade)	.00	.03	.02	.03	.02
	[.03]	[.16]	[.16]	[.16]	[.14]
Observations	72,561	72,955	74,769	79,663	299,948
Exit if fail at least one subject initially	.00	.09	.09	.12	.07
	[.05]	[.28]	[.29]	[.32]	[.26]
Observations	21,720	22,780	19,567	17,380	81,447

Panel B. Student Characteristics by subgroups

Variables	Female	White	Black	Hispanic	Economically Disadvantaged	Limited English Proficiency
Average Math Score (adjusted)	17.42 [41.33]	28.15 [37.38]	-9.52 [47.29]	-2.39 [44.68]	-5.19 [45.12]	-20.06 [45.05]
Average LAL Score (adjusted)	28.24 [35.83]	31.19 [31.44]	5.63 [45.50]	5.22 [44.78]	3.36 [44.10]	-33.37 [43.96]
Initial Passing Rate: Math	.73 [.44]	.85 [.36]	.44 [.50]	.52 [.50]	.49 [.50]	.28 [.45]
Initial Passing Rate: LAL	.89 [.31]	.93 [.26]	.71 [.45]	.7 [.46]	.68 [.47]	.22 [.42]
Exit (11th grade)	.02 [.13]	.01 [.10]	.05 [.22]	.04 [.20]	.04 [.20]	.07 [.26]
Observations	152,722	197,203	38,982	36,863	39,514	9,000
Exit if fail at least one subject initially	.07 [.25]	.06 [.23]	.09 [.28]	.08 [.28]	.08 [.27]	.08 [.28]
Observations	43,215	34,138	22,910	19,166	22,027	7,665

Table 1 Panel C. Indicators for Retest

Variables	Cohort1		Cohort 2		Cohort 3		Cohort 4		All cohorts	
	Math	LAL	Math	LAL	Math	LAL	Math	LAL	Math	LAL
Passing rate	.30	.43	.31	.59	.22	.45	.22	.37	.27	.46
	[.46]	[.50]	[.46]	[.49]	[.41]	[.50]	[.41]	[.48]	[.44]	[.50]
Observations	14,853	6,865	16,098	6,951	13,474	5,901	10,810	5,977	55,235	25,694
Exit if fail the retest	.01	.02	.04	.08	.04	.08	.05	.07	.04	.06
	[.12]	[.15]	[.19]	[.28]	[.21]	[.26]	[.22]	[.26]	[.19]	[.24]
Observations	10,456	3,895	11,087	2,866	10,558	3,270	8348	3,782	40,539	13,813

Variables	Female		White		Black		Hispanic		Economically Disadvantaged		Limited English Proficiency	
	Math	LAL	Math	LAL	Math	LAL	Math	LAL	Math	LAL	Math	LAL
Passing rate	.30	.44	.30	.68	.31	.42	.22	.34	.22	.35	.12	.16
	[.46]	[.50]	[.46]	[.47]	[.46]	[.49]	[.41]	[.47]	[.41]	[.48]	[.32]	[.37]
Observations	14,853	11,604	14,853	7,244	16,098	7,623	13,474	8,376	10,810	9,575	5,339	5,736
Exit if fail the retest	.03	.06	.04	.07	.04	.06	.04	.06	.04	.05	.04	.04
	[.17]	[.23]	[.19]	[.26]	[.18]	[.24]	[.20]	[.24]	[.20]	[.23]	[.21]	[.21]
Observations	24,420	6,456	12,481	2,336	14,375	4,408	11,454	5,554	13,571	6,191	4,720	4,790

Note: Panel A. Column 1,2,3,4 contains 11th graders in general education program in a public high school who took HSPA the first time in 2002, 2003, 2004, 2005 respectively. Sample in Column 5 is the sum of students from Column 1 to 4. Panel B are subsets of students from sample of Panel A's Column 5. Panel C includes students who didn't pass both subjects and enrolled in the 12th grade of The average math and LAL score are adjusted by subtracting the cut-off level 200 from the actual scale score. These are scores of the initial test administered in high school junior year. Standard deviations are in square brackets.

TABLE 2 Regression Discontinuity Estimates of Exiting from High School after Initial Attempt in HSPA

					Bandwidth			Cohort			
	All		All		+/-30	+/-40	+/-50	1	2	3	4
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Exit after failure†					.048(.214)	.055(.227)	.058(.234)				
Specification	cubic	cubic	quartic	quartic	linear	linear	linear	cubic	cubic	cubic	cubic
Math Initial Test	-.013***	-.011***	-.013***	-.010***	-.012***	-.010***	-.009***	-.000	-.015***	-.013***	-.017***
	[.002]	[.001]	[.002]	[.001]	[.001]	[.001]	[.001]	[.000]	[.003]	[.002]	[.004]
Observations	299,948	299,948	299,948	299,948	138,633	180,748	222,672	72,561	72,955	74,769	79,663
					Bandwidth			**Cohort**			
	All		All		+/-25	+30/-40	+/-50	1	2	3	4
Exit after failure†					.061(.239)	.068(.252)	.070(256)				
Specification	cubic	cubic	quartic	quartic	linear	linear	linear	cubic	cubic	cubic	cubic
LAL Initial Test	-.006***	-.005***	-.006***	-.005***	-.007***	-.005***	-.003***	-.000	-.007***	-.010***	-.004***
	[.002]	[.002]	[.002]	[.002]	[.001]	[.001]	[.001]	[.000]	[.003]	[.002]	[.002]
Observations	299,948	299,948	299,948	299,948	94,478	128,760	242,510	72,561	72,955	74,769	79,663
Controls	N	Y	N	Y	Y	Y	Y	Y	Y	Y	Y

Note: * significant at 10% ** significant at 5%; *** significant at 1% † Each cell is the average exit rate for students who score under the passing level within each bandwidth with standard deviation in parenthesis. Regressions on students who took the HSPA tests first time in 11th grade Spring semester, and without IEP exemption of passing. Students who stayed in the school less than one year when that test was administered were excluded. Dependent variable is a binary variable indicating whether the student exits from high school the semester following the initial test. Each cell represents a separate estimate of the discontinuity in failing the subject of testing listed on the first column and the specific functional form. Test scores are adjusted by subtracting the passing score 200 from the original scale score. Robust standard errors clustered on adjusted score level are in parentheses. All models contain a linear interaction term of test score and the passing status of the subject, allowing slopes to differ on both sides of the HSPA cutoff. Controls include dummies for female, special education students, economically disadvantaged students, LEP students, age, white, black, Hispanic, socioeconomic status of citizens in the district.

TABLE 3 RD Estimates of Exiting after Second Attempt in HSPA

	All Cohorts				Bandwidth		Cohort			
					-/+25	-/+30	1	2	3	4
Exit after failure†					.019(.136)	.021(.144)				
Specification	quadratic	quadratic	cubic	cubic	quadratic	cubic	quadratic	quadratic	quadratic	quadratic
Math 1st Retest	-.015***	-.013***	-.016***	-.015***	-.015***	-.015***	-.004**	-.011***	-.018***	-.024***
	[.002]	[.001]	[.003]	[.003]	[.002]	[.004]	[.002]	[.003]	[.004]	[.006]
Observations	55,235	55,235	55,235	55,235	30,496	41,499	14,853	16,089	13,474	10,810
Exit after failure†					.030(.171)	.031(.174)				
LAL 1st Retest	-.007	-.007**	-.012*	-.011**	-.016***	-.011***	-.005*	-.005	-.005	-.017**
	[.005]	[.003]	[.007]	[.005]	[.005]	[.004]	[.003]	[.007]	[.007]	[.004]
Observations	25,694	25,694	25,694	25,694	13,980	18,188	6,865	6,951	5,901	5,977
Controls	N	Y	N	Y	Y	Y	Y	Y	Y	Y

Note: * significant at 10% ** significant at 5%; *** significant at 1%. † Each cell is the average exit rate for students who score under the passing level within each bandwidth with standard deviation in parenthesis. Standard errors in brackets robust to heteroskedasticity and correlation within clusters. Dependent variable is a binary variable indicating whether the student exits from high school the semester following the second test. Each cell represents a separate estimate of the discontinuity in failing the subject of testing listed in the left hand column. Details see Note under Table 2.

TABLE 4 RD Estimates of Exiting from High School for Subgroups (Initial Attempt)

	Female		Limited English Proficiency		White	
Specification	cubic	quartic	cubic	quartic	cubic	quartic
Math	-.007***	-.007***	-.016	-.023	-.005***	-.004***
Initial Test	[.001]	[.001]	[.010]	[.016]	[.001]	[.001]
LAL	-.001	-.002**	-.041*	-.030	-.005***	-.004**
Initial Test	[.001]	[.001]	[.021]	[.022]	[.001]	[.001]
Observations	152,722	152,722	9,000	9,000	197,203	197,203
Controls	Y	Y	Y	Y	Y	Y

	Hispanic		Black		Economically Disadvantaged	
Specification	cubic	quartic	cubic	quartic	cubic	quartic
Math	-.013***	-.015***	-.014***	-.015***	-.013***	-.016***
Initial Test	[.002]	[.002]	[.003]	[.003]	[.002]	[.003]
LAL	-.008***	-.010***	.000	-.003	-.006***	-.006***
Initial Test	[.002]	[.002]	[.002]	[.002]	[.001]	[.002]
Observations	36,863	36,863	38,982	38,982	39,514	39,514
Controls	Y	Y	Y	Y	Y	Y

Note: * significant at 10% ** significant at 5%; *** significant at 1%. Standard errors in brackets robust to heteroskedasticity and correlation within clusters. Regressions on the specified subgroup students who took the HSPA tests first time in 11th grade Spring semester (Table 4) and without IEP exemption of passing. Students who stayed in the school less than one year when that test was administered were excluded. Dependent variable is a binary variable indicating whether the student exits from high school the semester following the first test. Each cell represents a separate estimate of the discontinuity in failing the subject of testing listed in the left hand column. Details see footnotes under Table 2.

TABLE 5 RD Estimates of Exiting from High School for Subgroups (Retest)

	Female		Limited Engligh Proficiency		White	
Specification	quadratic	cubic	quadratic	cubic	quadratic	cubic
Math	-.012***	-.012***	-.006	-.008	-.014***	-.014***
1st Retest	[.002]	[.003]	[.006]	[.010]	[.003]	[.003]
	32,810	32,810	5,339	5,339	20,773	20,773
LAL	-.010**	-.014**	-.004	-.014	-.008	-.009*
1st Retest	[.004]	[.006]	[.005]	[.014]	[.005]	[.005]
Observations	11,604	11,604	5,736	5,736	7,244	7,244
Controls	Y	Y	Y	Y	Y	Y

	Hispanic		Black		Economically Disadvantaged	
Specification	quadratic	cubic	quadratic	cubic	quadratic	cubic
Math	-.013***	-.018**	-.009***	-.009	-.012***	-.014**
1st Retest	[.003]	[.008]	[.002]	[.005]	[.003]	[.009]
	14,147	14,147	17,067	17,067	16,305	16,305
LAL	-.002	-.027	-.012*	-.015	-.000	-.007
1^{st} Retest	[.002]	[.016]	[.007]	[.011]	[.004]	[.007]
Observations	8,376	8,376	7,623	7,623	9,575	9,575
Controls	Y	Y	Y	Y	Y	Y

Note: * significant at 10% ** significant at 5%; *** significant at 1%. Standard errors in brackets robust to heteroskedasticity and correlation within clusters. Regressions on the specified subgroup who took HSPA retest in 12^{th} grade Fall semester and without IEP exemption of passing. Students who stayed in the school less than one year when that test was administered were excluded. Dependent variable is a binary variable indicating whether the student exits from high school the semester following second test. Each cell represents a separate estimate of the discontinuity in failing the subject of testing listed in the left hand column. Details see notes under Table 2.

TABLE 6 Effects of Failing At Least One Subject in HSPA

	All		Cohort 1		Cohort 2		Cohort 3		Cohort 4	
	1st Test	Retest	1st Test	Retest	1st Test	Retest	1st Test	Retest	1st Test	Retest
Specification	quadratic	linear	quadratic	linear	quadratic	linear	quadratic	linear	quadratic	linear
Math Test	-.007	-.010**	-.000	-.002	.003	-.013	-.026*	-.015	-.002	-.007
	[.007]	[.004]	[.002]	[.006]	[.014]	[.014]	[.015]	[.011]	[.009]	[.018]
Observations	39,325	11,921	10,649	3,439	10,589	2,590	8,940	2,854	9,147	3,038
LAL Test	-.003	-.007*	-.000	-.007*	-.011*	-.006	-.008	-.003	.013	-.012
	[.006]	[.003]	[.001]	[.003]	[.007]	[.008]	[.005]	[.009]	[.009]	[.008]
Observations	73,459	16,894	19,729	4,486	20,874	4,621	17,671	4,011	15,185	3,776
Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

Note: * significant at 10% ** significant at 5%; *** significant at 1% Standard errors in brackets robust to heteroskedasticity and correlation within clusters. Dependent variable is a binary variable indicating whether the student exits from high school the semester following the initial or second test. Each cell represents a separate estimate of the discontinuity in failing the subject of testing listed in the left hand column conditional on failing the other subject. Functional form of the test score is specified in row 3. Details see Note in Table 2.

Appendix I: Data construction and variable description

a. Data construction

The raw data is provided by NJDOE which contain student characteristics and test scores in math and Language Art Literacy for all test takers in each test administration from March 2002 to March 2006.

Table: HSPA Administration by Academic Cohort

Cohort	Spring 2002	Fall 2002	Spring 2003	Fall 2003	Spring 2004	Fall 2004	Spring 2005	Fall 2005	Spring 2006
2003	X	X	X						
2004			X	X	X				
2005					X	X	X		
2006							X	X	X

As indicated from the table above, I trace students in each cohort who took HSEE from their 11th grade to the end of 12th grade. Information on all three tests (including two retests) will be combined if applicable. Data sets of three test administrations were merged using a unique student HSPA test id and school ID[1]. In order to ensure I correctly match students over three testing administration, I also compared the date of birth information. I then compare the sample size with the official publicized statistics on NJDOE website[2], which proves that the numbers are identical before I do further selections (or deletions) and merging.

I cleaned the data by cohort, and then merged the clean data sets of four cohorts together into a pooled sample. For each cohort, I deleted observations that did not have consistent date of birth (DOB) information in different test administrations. Students who are older than 21 years old and younger than 15 years old are also excluded. I took out all migrant students (27 in total from 2002 to 2006). Moreover, I excluded from the sample any students with missing or wrong coding on gender, grade, and ethnicity. The cleaned sample size for each cohort is 84,325; 103,838; 108,704 and 118,078 respectively. The pooled data therefore contains 414,945 students for all four cohorts.

I only include in my analysis 356,359 observations who took HSPA initially in the spring semester of 11th grade. I include the students who attended some professional courses but enrolled in a general education (GE) program in a public school. However, I do not consider students who are from vocational schools, private schools, and rural districts (18270 students). I further exclude 12,270 students who stayed in school for less than one year when the test was administered, as well as 24,613 special education (SE) students who got the IEP exemption in the first test and 1,258 SE students who got the IEP exemption in the retest. This leaves to a final sample of 299,948 observations.

The retest sample includes all student who failed at lest one subject initially and were enrolled in fall semester. I exclude all special education students (N= 9741) because the special education students do not seem to randomly distributed around the cut-off score from the RD validity test.

[1] Students who transferred to a different school within state during the test administrations are not included in my sample. The DFG code of school would be used in matching to increase accuracy.

[2] http://www.state.nj.us/education/assessment/hs.

b. Variables

Outcome variables:

Exiting from High School after Initial Attempt & *Exiting from High School after Second Test*
The "exit" status is indicated in the retest records. This is from the "Exit Student Roster" in NJ HSPA database, which is reported and updated by the school in which the student is enrolled, the beginning of every semester. If a student pass the exam in the first test and leave the school without continuing registration for the next semester/term in the school or other high schools in New Jersey, the school enrollment would be shown as "exit". So more precisely, the "exit" represents student's dropout episode occurred between the two test administrations, for example, exiting from high school after the first exam, means student dropout before starting the 12^{th} grade fall semester. Similarly, "exit" from high school after the second test, means students did not enroll in school for the spring semester of 12^{th} grade[3].

Other variables:

Female: Dummy variable for female students. Recode from original variable "sex".
Special education: Dummy variable. Students who were coded in any special education program received a value of one.
LEP: Dummy variable. It equals to one if the student is identified as limited English proficiency.
ED: Dummy variable. It equals to one for economically disadvantaged students.
Age: Age in the first test is calculated by DOB and March 1^{st} of the testing year and for age in the retest, I used the DOB and September 1^{st} of the testing administration year for calculation.
Older students: Older is defined as students who are equal or older than 20 years old at the initial test. The assumption of including this dummy variable is that these students at the fringe of being "required" to leave regular high school (for example, to enroll in an adult high school after turning 21) within one year, so they might have more pressure of passing the exam in order to graduate on time.
White/Black/Hispanic: Dummy variables for white, black or Hispanic students. The omitted category is Asian and other race/ethnicity.
SES1/ SES2/SES3: Socioeconomic status grouping. The original data from NJDOE contains a variable called DFG (District Factor Group): this is an indicator of socioeconomics status of citizens in each district which was developed by NJDOE. It was calculated using several demographic variables from 2000 United States Census including a) percent of adult residents who failed to complete high school; b) percent of adult residents who attended college; c) occupational status of adult household members ; d) unemployment rate ; e) poverty rate ; and f) median household income. There are eight categories of DFG code A, B, CD, DE, FG, GH, I, J, where A stands for the lowest socioeconomic districts and J stands for the highest socioeconomic districts. SE1 is a recoding for A & B; SE2 is a recoding fro CD & DE ; and SE3 is a recoding for FG & GH. The omitted SE group is I & J.

[3] Any within-state transfer will be captured by the data base; however, out-of-state transfer would be treated as "exit" though the percentage of students transferred out from New Jersey high school is very small and would not induce a biased estimate in this paper.

Appendix Table A-I: State High School Exit Exam Basic Features

State	HSEE Type	First Class Required HSEE for Graduation	Test on 9th/10th Grade (Yes=1; No=0)	Subjects Tested
Alabama	SBE	2001	0*	R,L,M, SC,SS
Alaska	MCE	2004	1	R,W,M
Arizona	SBE	2006	1	R,W,M
Arkansas				
California	SBE	2006	1	ELA, M
Colorado				
Connecticut				
Delaware				
D.C.				
Florida	SBE	2003	1*	R,M
Georgia	SBE	1994	0*	ELA,W,M,SC,SS
Hawaii				
Idaho	SBE	2006	1	R,L,M, SC
Illinois				
Indiana	SBE	2000	1	ELA, M
Iowa				
Kansas				
Kentucky				
Louisiana	SBE	2003	0*	ELA, M, SC, SS
Maine				
Maryland	EOC	2009	1*	note(5)
Massachusetts	SBE	2003	1	ELA, M, SC(2010)
Michigan				
Minnesota	MCE	2000/2010(SBE)	1	R,W,M
Mississippi	EOC	2003	0*	notes(6)
Missouri				
Montana				
Nebraska				
Nevada	SBE	2003	1*	R,W,M, SC(2008)
New Hampshire				
New Jersey	SBE	2003	0*	LAL, M, S
New Mexico	MCE	1990	1	R,L,C, SC, SS
New York	EOC	2000	0*	ELA, M, SC, SS
North Carolina	SBE	1982/2001	1	note(8)
Ohio	SBE	2007	1**	R, W, M, SS, SC
Oklahoma	EOC	2012	1	note (9)
Oregon				
Pennsylvania				
Rhode Island				
South Carolina	SBE	2006	1**	ELA, M, SC(2010), H
South Dakota				
Tennessee	EOC	1992/2005	0**	(10)
Texas	SBE	2005	0*	ELA, M,SC, SS
Utah	MCE	2006	1	-
Vermont				
Virginia	EOC	2004	0*	(11)
Washington	SBE	2008	1	R, W, M, SC (2010)
West Virginia				
Wisconsin				
Wyoming				

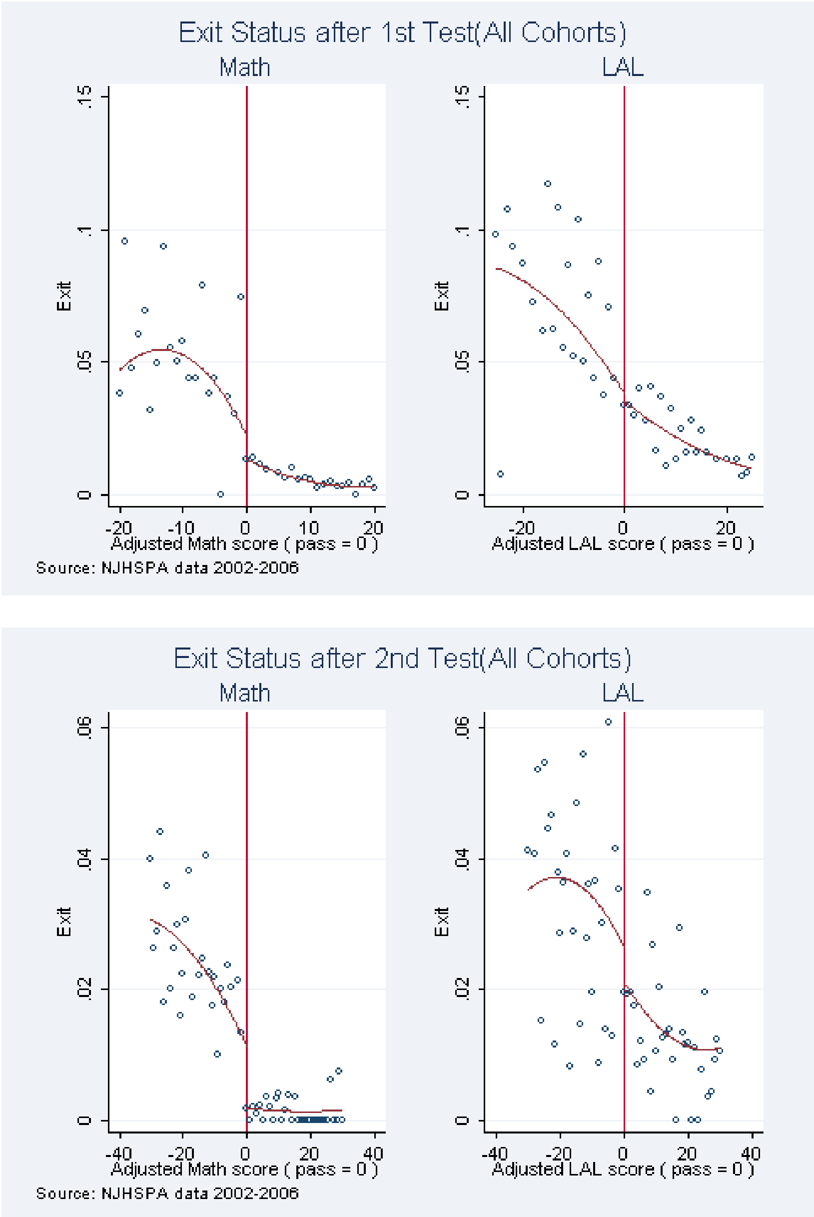
TABLEA-II Regression Discontinuity Estimates on Pre-Determined Student Characteristics

Variables	**Math**		**Math (Failed LAL)**		**LAL**		**LAL (Failed Math)**	
	1st Test	**Retest**	**1st Test**	**Retest**	**1st Test**	**Retest**	**1st Test**	**Retest**
Female	-.004	.001	-.010	.027	-.025***	-.047***	.001	-.016
	(.008)	(.008)	(.017)	(.027)	(.008)	(.011)	(.012)	(.015)
White	-.010	-.049***	-.013	.003	-.000	.015	.005	.013
	(.011)	(.009)	(.017)	(.024)	(.011)	(.017)	(.011)	(.017)
Black	-.007	.006	.003	-.034*	-.008*	.000	-.018**	-.025
	(.006)	(.005)	(.005)	(.020)	(.004)	(.015)	(.009)	(.020)
Hispanic	.005	.018***	.011	.012	.012*	.023***	.009	.025**
	(.009)	(.006)	(.010)	(.021)	(.007)	(.009)	(.007)	(.011)
Economically Disadvantaged	.009*	.012**	.026***	-.020	.002	.005	-.006	.004
	(.005)	(.006)	(.008)	(.025)	(.006)	(.011)	(.008)	(.013)
Limited English Proficiency	.003	.012***	-.007	-.017	.006**	.014**	.004***	.007*
	(.003)	(.003)	(.007)	(.030)	(.003)	(.007)	(.001)	(.004)
Older Students	.001	.000	.001	.000	-.002	-.000	.000	.000
	(.001)	(.000)	(.001)	(.000)	(.002)	(.000)	(.000)	(.000)
Special Education Student	.011**	-	.011	-	.006*	-	-.006	-
	(.004)	-	(.007)	-	(.003)	-	(.006)	-
SES 1	.015	.038***	.014	.017	-.009	.006	-.017	-.004
	(.010)	(.008)	(.013)	(.025)	(.007)	(.014)	(.010)	(.016)
SES 2	-.010	-.026***	-.039***	.008	.032***	.040***	.006	.019
	(.007)	(.009)	(.015)	(.021)	(.007)	(.008)	(.008)	(.013)
SES 3	-.005	-.017**	.011	.005	.012	-.001	.009	-.008
	(.004)	(.007)	(.014)	(.021)	(.007)	(.011)	(.007)	(.012)
Cohort 1	-.029	.004	-.027	-.031	-.012	-.032	-.023	-.031
	(.046)	(.069)	(.066)	(.091)	(.067)	(.111)	(.073)	(.124)
Cohort 2	-.031	-.020	-.009	.008	-.014	.047	-.009	.054
	(.072)	(.079)	(.072)	(.022)	(.073)	(.083)	(.082)	(.089)
Cohort 3	.063	.028	.085	-.025	.062	.002	.043	.013
	(.084)	(.067)	(.108)	(.094)	(.123)	(.078)	(.109)	(.089)
Observations								
N	299,948	55,235	39,325	11,921	299,948	25,694	73,459	16,894

Notes for Table A-I :
(1) Information retreived and complied from Center on Education Policy (2004, 2005, 2006,2007). www.cep-dc.org
(2)EOC= End-of-Course Exam MCE=Minimum Competency Exam SBE=Standards-based Exam (3) C= composition L=Language M=Math SC=Science ELA/LAL=English/Language literacy arts R= Reading SS=Social Studies W=Writing H= U.S. History (3) Maryland: It has HSEE but it does not withhold the diplomas for HS students who didn't pass the test.
(4) * Prior HSEE exists ** Prior HSEE is an MCE (5) Eglish II Algebra/data analysis, biology, government
(6) English II (with writing), Algebra I, Biology I, U.S. History; (7) New York also test language other than English; (8) Reading, math, computer skills; (9) English II, III, Algebra I, Algebra II, geometry, Biology I, U.S. history; (10) same as in note (6) but without history; (11) English, Algebra I & II, geometry, biology, earth science, chemistry, world history, Viginia and U.S. history, world geography

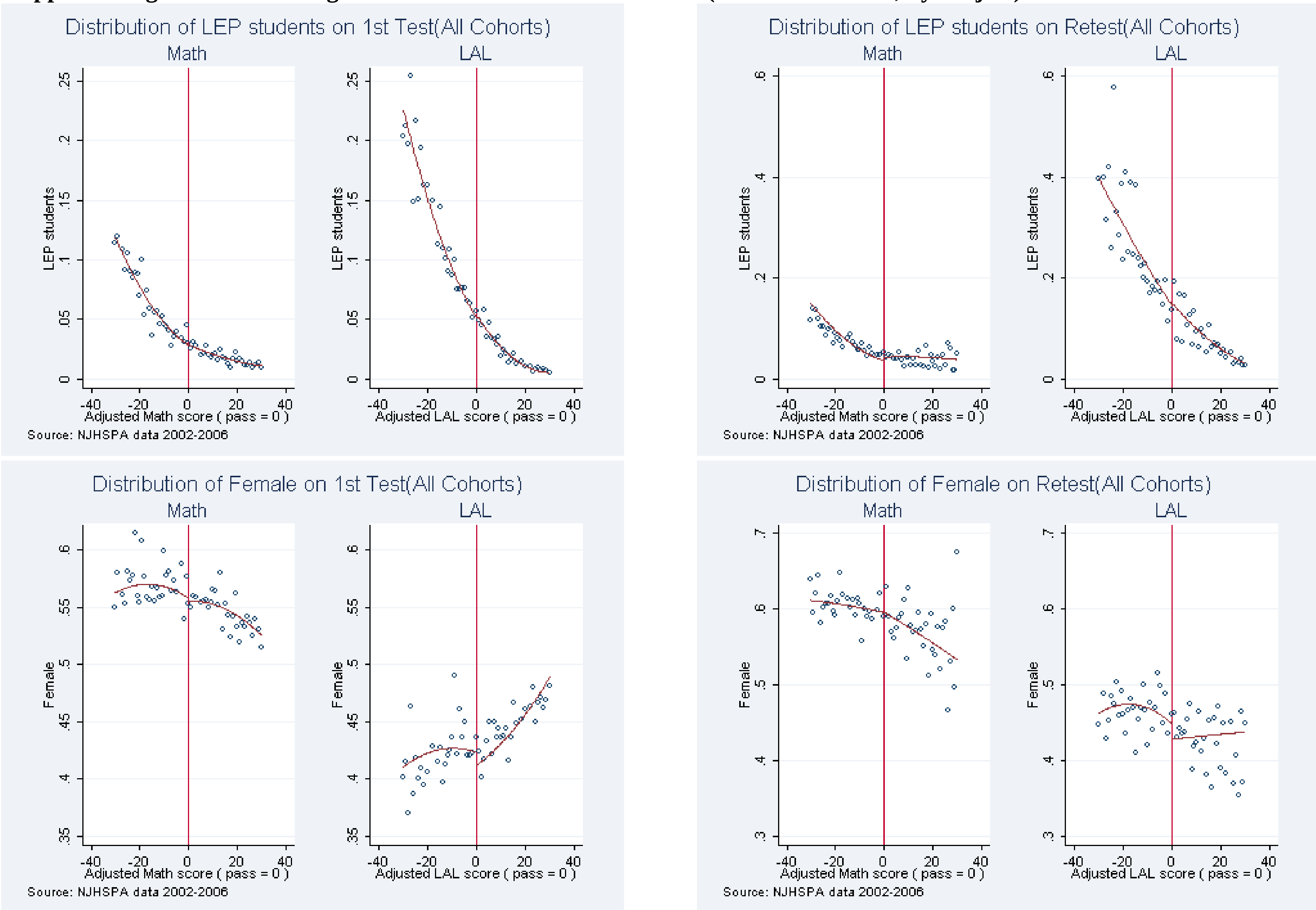
Notes for Table A-II:
(1) Data source: New Jersey High School Proficiency Assessment (HSPA) Data base 2002-2006. (2) * p<.1 ** p<.05 *** p<.01. (3) All scores are adjusted relative to the passing cut-off of HSPA test.(4) All estimates for the 1st test are marginal effect of the dependent variables specified in the first column at the pass cut-off in the probit regression based on the fourth polynomial form of the specific test score using the selected samples. (5) All estimates for the retest are marginal effect of the dependent variables specified in the first column at the pass cut-off in the probit regression based on the quadratic form of the specific test score with the dummy indicating passing the test as well as other student. (7) Standard errors are clustered at the score level. (6) Special education students with exemption to pass the HSPA are excluded from the analysis of the initial test; all special education students are excluded in the analysis of the retest. (7) See Appendix I for details on variable descriptions and sample selection.

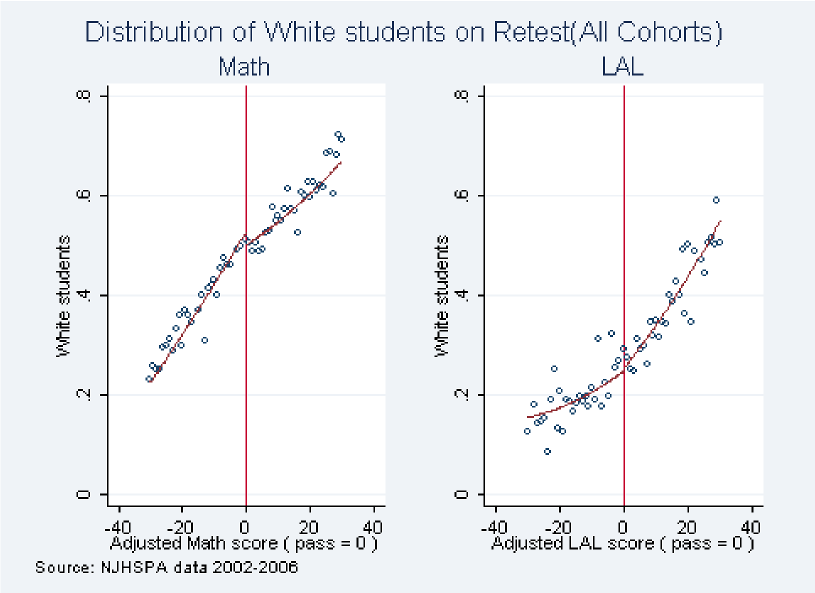
Figure A-I. The Relationship between 1st and 2nd HSPA Test Score and Exiting High School by Testing Subject for four cohorts (2002-2005)

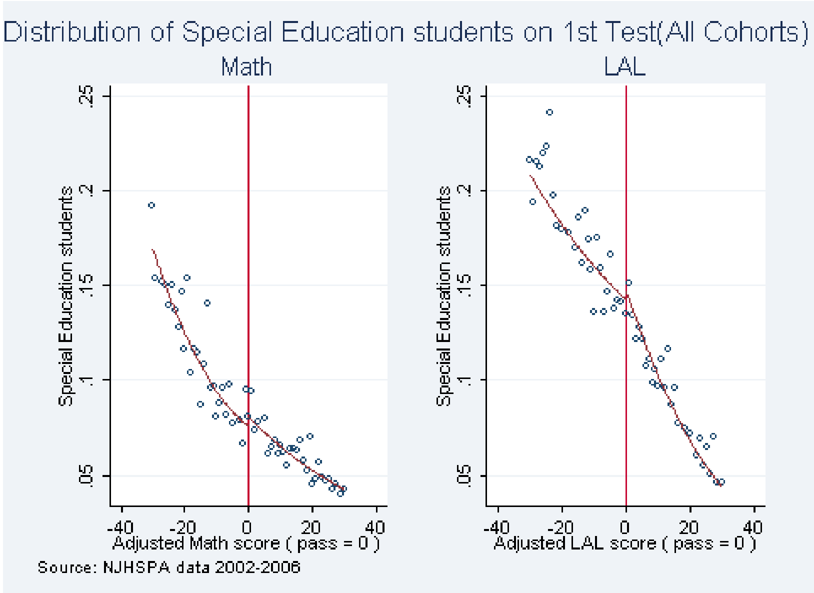


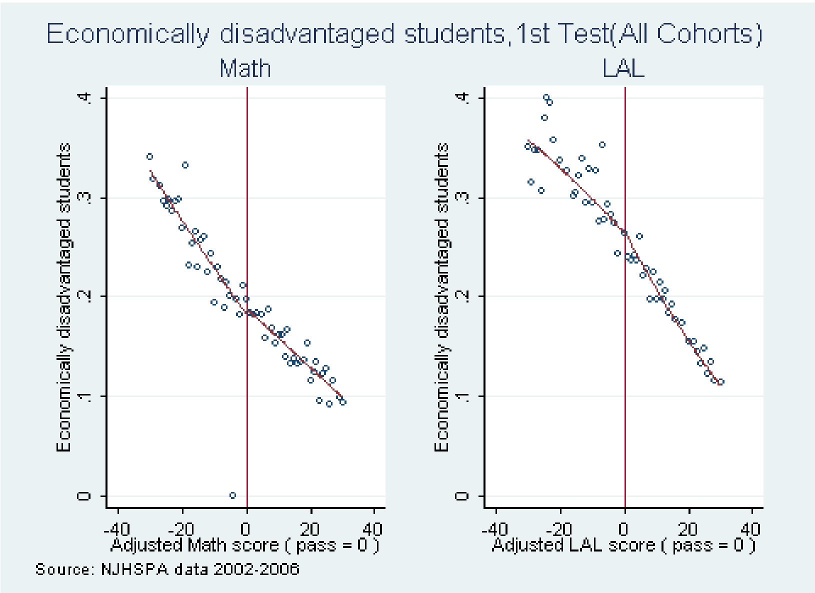
Note: The left panel shows the probability of exiting (Y-axis) from high school along the adjusted-math score (X-axis); the right panel shows the probability of exiting (Y-axis) from high school along the adjusted-LAL score (X-axis). The circles reflect the actual average probability of exiting at each score point. The solid line represents the predicted exit probability from a fourth order polynomial function of the initial test score (upper graph) or a cubic form for the retest score (lower graph), along with an interaction between the passing dummy and a linear term of the test score.

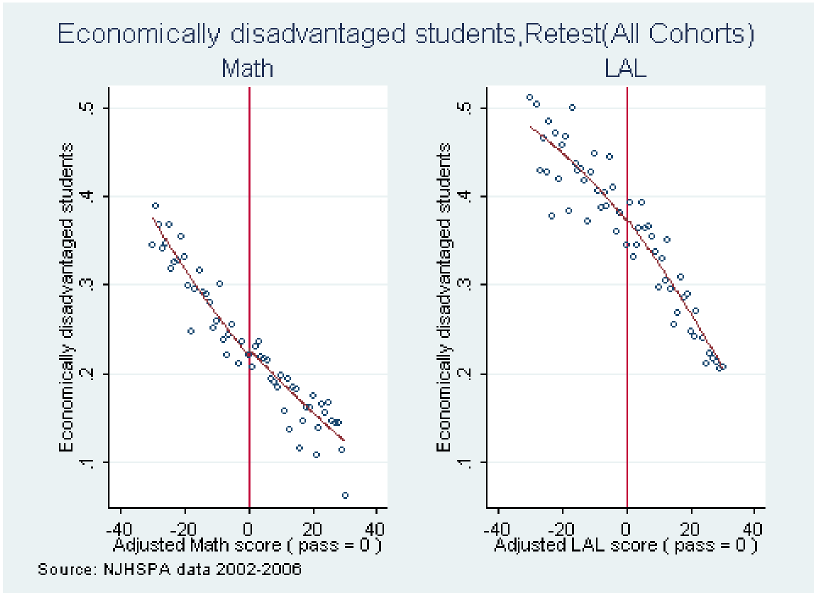
Appendix Figure A-II. Plotting of selected baseline characteristics (Initial and Retest, by subject)




Distribution of White students on Retest(All Cohorts)
Math
LAL
White students
Adjusted Math score (pass = 0)
Adjusted LAL score (pass = 0)
Source: NJHSPA data 2002-2006


Distribution of Special Education students on 1st Test(All Cohorts)
Math
LAL
Special Education students
Adjusted Math score (pass = 0)
Adjusted LAL score (pass = 0)
Source: NJHSPA data 2002-2006


Economically disadvantaged students, 1st Test(All Cohorts)
Math
LAL
Economically disadvantaged students
Adjusted Math score (pass = 0)
Adjusted LAL score (pass = 0)
Source: NJHSPA data 2002-2006


Economically disadvantaged students,Retest(All Cohorts)
Math
LAL
Economically disadvantaged students
Adjusted Math score (pass = 0)
Adjusted LAL score (pass = 0)
Source: NJHSPA data 2002-2006

Note:
In each graph, the left panel shows the percentage of each specified baseline characteristic (Y-axis) along the adjusted-math score (X-axis); the right panel shows the percentage of each specified baseline characteristic (Y-axis) along the adjusted-LAL score (X-axis). The circles reflect the actual average proportion of the outcome at each score point. The solid line represents the predicted values from a fourth order polynomial function in the initial test score (upper graph) or a cubic function in the retest score (lower graph) fitted separately for students above and below the passing cut-off.

Reference:

Angrist, J, and Lavy, V. (1999). 'Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement.' Quarterly Journal of Economics, 114(2), pp.533-75.

Belfield, Clive, and Levin, H.M. (2007). 'The Economic Losses from High School Dropouts in California.' California Dropout Research Project Report 1.

Bishop, J. H. (1998). 'The effect of Curriculum-Based External Exit Exam Systems on Student Achievement.' The Journal of Economic Education, 29(2), pp. 171- 82

Bishop, J. H. (2005). 'High School Exit Examinations, pp. When Do Learning Effects Generalize?' In *Uses and Misuses of Data in Accountability Testing, National Society for the Study of Education Yearbook*, ed. Joan Herman and Edward Haertel, Part I. Chicago, pp. University of Chicago Press.

Card, D., and Lee, D. (2008). 'Regression Discontinuity Inference with Specification Error', *Journal of Econometrics,* 142(2), pp. 655-74.

Carnoy, M., and Loeb, S. (2002). 'Does external accountability affect student outcomes? A cross-state analysis.' Educational Evaluation and Policy Analysis, 24(4), 305-31.

Chay, K. Y., McEwan, P. J. and Urquiola, M. (2005). 'The Central Role of Noise in Evaluating Interventions That Use Test Scores to Rank Schools.' American Economic Review, 95(4), pp. 1237–58

Center on Education Policy (CEP). (2004). State High School Exit Exams, pp. A Measuring Reform. Washington, DC, pp. Author.

Center on Education Policy (CEP). (2006). State High School Exit Exams, pp. A Challenging Year. Washington, DC, pp. Author.

Center on Education Policy (CEP). (2007). State High School Exit Exams, pp. Working to Raise Test Scores. Washington, DC, pp. Author.

Cornell, D. G., Krosnick, J. A. and Chang, L.C. (2006). 'Student Reactions to Being Wrongly Informed of Failing a High-Stakes Test.' Educational Policy, *20*(5). pp.718-51.

Cullen, J. B., and Reback, R. (2006). 'Tinkering Toward Accolades, pp. School Gaming Under A Performance Accountability System.' National Bureau of Economic Research Working Paper 12286.

Dee, T.S. (2002). 'the 'First Wave' of Accountability.' in P. Petersen and M. West (eds), *No Child Left Behind? The Politics and Practice of Accountability,* Brookings Institution Press.

Dee, T. S., and Jacob, B.A. (Forthcoming). 'Do High School Exit Exam Influence Education Attainment or Labor Market Performance?' In *Will No Child Left Behind Help Close the Poverty Gap?*, ed. Adam Gamoran, Washington, D.C., pp. Brookings University Press.

Figlio, D., and Getzler, L. S. (2002). 'Accountability, Ability and Disability, pp. Gaming the System.' National Bureau of Economic Research Working Paper 9307.

Fine, M., Pappas, L., Karp, S., Hirsch, L., Sadovnik, A., Keeton, A. and Bennett, M. (2007). 'New Jersey's Special Assessment, pp. Loophole or Lifeline? A Policy Brief.' Education Law Center.

Fuller, E. J., and Johnson, J.F.. (2001). 'Can State Accountability Systems Drive Improvements in School Performance for Children of Color and Children from Low-Income Homes?' Education and Urban Society, 33 (3), pp. 260-83.

Garcia, P., and Gopal, M. (2003). 'The Relationship to Achievement on the California High School Exit Exam for Language Minority Students.' NABE Journal of Research and Practice, 1, pp. 123-37.

Griffin, B. W., and Heidorn, M.H. (1996). 'An Examination of the Relationship Between

Minimum Competency Test Performance and Dropping Out of High School.' Educational Evaluation and Policy Analysis. 18, pp. 243-52.
Greene, J. P., and Winters, M.A. (2004). 'Pushed Out or Pushed Up? Exit Exams and Dropout Rates in Public High Schools.' Manhattan Institute Education Working Paper 5.
Hanushek, E. A., and Raymond, M. (2004). 'The effect of school accountability systems on the level and distribution of student achievement.' Journal of the European Economic Association, 2(2-3), pp.406-15.
Hanushek, E. A., and Raymond, M. (2005). 'Does school accountability lead to improved student performance?' Journal of Policy Analysis and Management, 24(2), pp. 297-327.
Harris, D. N., and .Herrington, C.D. (2006). 'Accountability, Standards, and the Growing Achievement Gap, pp. Lessons from the Past Half-Century.' American Journal of Education, 112 (2), pp.209-38
Hoffman, J. V., Assaf, L.C. and Paris, S.G. (2001). 'High-stakes testing in reading, pp. Today in Texas, tomorrow?' The Reading Teacher, 54(5), pp.482-92.
Imbens, G, and Lemieux, T. (2008). 'Regression Discontinuity Designs, pp. A Guide to Practice.' Journal of Econometrics, 142(2), pp. 615-35.
Jacob, B. A. (2001). 'Getting tough? The Impact of High School Graduation Exams.' Educational Evaluation and Policy Analysis, 23(2), pp. 99-122.
Jacob, B. A., and Lefgren, L.(2004). 'Remedial Education and Student Achievement , pp. A Regression-Discontinuity Analysis.' Review of Economics and Statistics, 86(1), pp. 226-44.
Jacob, Brian A., and Lars Lefgren. (2007). 'The Effect of Grade Retention on High School Completion.' National Bureau of Economic Research Working Paper 13514.
Jacob, B. A., and Levitt, S.D. (2003). 'Rotten Apples, pp. An Investigation of the Prevalence and Predictors of Teacher Cheating.' Quarterly Journal of Economics, 118(3), pp.843-77.
Jones, M. G., Jones, B.D., Hardin,B., Chapman, L., Yarbrough, T., and Davis, M. (1999). 'The Impact of High-Stakes Testing on Teachers and Students in North Carolina.' Phi Delta Kappan, 81(3), pp.199-203.
Lazear, E. P. (2004). 'Speeding, Tax Fraud, and Teaching to the Test.' National Bureau of Economic Research Working paper 10932.
Lemieux, T., and Milligan, K. (2008). 'Incentive effects of social assistance, pp. A regression Discontinuity Approach.' Journal of Econometrics, 127(2), pp. 807-28.
Loeb, S., and Strunk, K. (2003). 'The contribution of administrative and experimental data to education policy research.' National Tax Journal, 56 (2), pp.415-38.
Kane, T. J. (2003). 'A Quasi-experimental Estimate of the Impact of Financial Aid on College-going.' National Bureau of Economic Research Working Paper 9703.
Mackey, P. E. (2006). *New Jersey's Public Schools, pp. A Biennial Report for the People of New Jersey, (2006)-07 edition.* New Brunswick, NJ, pp. Public Education Institute.
Martorell, F. (2004). 'Do High School Graduation Exams Matter? Evaluating the Effects of Exit Exam Performance on Student Outcomes.' PhD Dissertation Chapter, University of California, Berkeley.
McCrary, J., and Royer, H. (2005). 'The Effect of Maternal Education on Fertility and Infant Health, pp. Evidence form School Entry Policies Using Exact Date of Birth.' http, pp.//sitemaker.umich.edu/hroyer/files/jmp.pdf.
McEwan, P. J., and Shapiro, J.S. (2008). 'The Benefits of Delayed Primary School Enrollment. Discontinuity Estimates Using Exact Birth Rates.' Journal of Human Resources, 43(1), pp.1-29

Muller, C. (1998). ‘The Minimum Competency Exam Requirement, Teachers’ and Students’ Expectations and Academic Performance.’ Social Psychology of Education, 2 (2), pp.199-216.
Muller, C, and Schiller, K.S. (2000). ‘Leveling the Paying Field? Students’ Educational Attainment and States’ performance Testing.’ Sociology of Education, 73(3), pp.196-218.
New Jersey Department of Education. (2007). ‘March 2007 and October 2007 High School Proficiency Assessment (HSPA) Cycle I and Cycle II Score Interpretation Manual.’ NJDOE PTM 1506.11.
Papay, J., Murnane, R.J., and Willett, J.B. (2008). ‘The Consequences of High School Exit Examinations for Struggling Low-Income Urban Students, pp. Evidence from Massachusetts.’ National Bureau of Economic Research Working Paper 14186.
Reardon, S. F., Arshan, N., Atteberry, A. and Kurlaender, M. (2008). ‘High Stakes, No Effects, pp. Effects of Failing the California High School Exit Exam.’ Paper presented at the Annual Meeting of the Association for Pubic Policy Analysis and Management, Los Angeles.
Reback, R. (2008). ‘Teaching to the Rating, pp. School Accountability and the Distribution of Student Achievement.’ Journal of Public Economics, 92(5-6), pp. 1394-1415.
Trochim, W. M. K. 1984. ‘Research Design for Program Evaluation, pp. The Regression Discontinuity Approach.’ Beverly Hills, CA, pp. Sage Publications.
van der Klaauw, W. (2002). ‘Estimating the Effect of Financial Aid Offers on College Enrollment, pp. A Regression-Discontinuity Approach.’ International Economic Review, 43(4), pp. 1249-87.
Warren, J. R., Jenkins, K.N. and Kulick, R.B. (2006). ‘High School Exit Examinations and State-level Completion and GED Pass Rates, pp. 1975-2002.’ Educational Evaluation and Policy Analysis, 28(2), pp.131-52.
Warren, J. R., and Edwards, M.R. (2005). ‘High School Exit Examinations and High School Completion, pp. Evidence From the Early 1990s.’ Educational Evaluation and Policy Analysis, 27(1), pp. 53-74.

CENTRE FOR ECONOMIC PERFORMANCE
Recent Discussion Papers

906	Andrew B. Bernard J. Bradford Jensen Stephen J. Redding Peter K. Schott	The Margins of US Trade
905	Gianluca Benigno Bianca De Paoli	On the International Dimension of Fiscal Policy
904	Stephen J. Redding	Economic Geography: A Review of the Theoretical and Empirical Literature
903	Andreas Georgiadis Alan Manning	Change and Continuity Among Minority Communities in Britain
902	Maria Bas	Trade, Technology Adoption and Wage Inequalities: Theory and Evidence
901	Holger Breinlich Chiara Criscuolo	Service Traders in the UK
900	Emanuel Ornelas John L. Turner	Protection and International Sourcing
899	Kosuke Aoki Takeshi Kimura	Central Bank's Two-Way Communication with the Public and Inflation Dynamics
898	Alan Manning Farzad Saidi	Understanding the Gender Pay Gap: What’s Competition Got to Do with It?
897	David M. Clark Richard Layard Rachel Smithies	Improving Access to Psychological Therapy: Initial Evaluation of the Two Demonstration Sites
896	Giorgio Barba Navaretti Riccardo Faini Alessandra Tucci	Does Family Control Affect Trade Performance? Evidence for Italian Firms
895	Jang Ping Thia	Why Capital Does Not Migrate to the South: A New Economic Geography Perspective
894	Kristian Behrens Frédéric Robert-Nicoud	Survival of the Fittest in Cities: Agglomeration, Selection and Polarisation
893	Sharon Belenzon Mark Schankerman	Motivation and Sorting in Open Source Software Innovation
892	Guy Michaels Ferdinand Rauch Stephen J. Redding	Urbanization and Structural Transformation

891	Nicholas Bloom Christos Genakos Ralf Martin Raffaella Sadun	Modern Management: Good for the Environment or Just Hot Air?
890	Paul Dolan Robert Metcalfe	Comparing willingness-to-pay and subjective well- being in the context of non-market goods
889	Alberto Galasso Mark Schankerman	Patent Thickets and the Market for Innovation: Evidence from Settlement of Patent Disputes
888	Raffaella Sadun	Does Planning Regulation Protect Independent Retailers?
887	Bernardo Guimaraes Kevin Sheedy	Sales and Monetary Policy
886	Andrew E. Clark David Masclet Marie-Claire Villeval	Effort and Comparison Income Experimental and Survey Evidence
885	Alex Bryson Richard B. Freeman	How Does Shared Capitalism Affect Economic Performance in the UK?
884	Paul Willman Rafael Gomez Alex Bryson	Trading Places: Employers, Unions and the Manufacture of Voice
883	Jang Ping Thia	The Impact of Trade on Aggregate Productivity and Welfare with Heterogeneous Firms and Business Cycle Uncertainty
882	Richard B. Freeman	When Workers Share in Profits: Effort and Responses to Shirking
881	Alex Bryson Michael White	Organizational Commitment: Do Workplace Practices Matter?
880	Mariano Bosch Marco Manacorda	Minimum Wages and Earnings Inequality in Urban Mexico. Revisiting the Evidence
879	Alejandro Cuñat Christian Fons-Rosen	Relative Factor Endowments and International Portfolio Choice
878	Marco Manacorda	The Cost of Grade Retention
877	Ralph Ossa	A ‘New Trade’ Theory of GATT/WTO Negotiations

The Centre for Economic Performance Publications Unit
Tel 020 7955 7284 Fax 020 7955 7595 Email info@cep.lse.ac.uk
Web site http://cep.lse.ac.uk