ENGLISH PROGRAMS IN OREGON: HELPING OR INHIBITING ENGLISH LANGUAGE LEARNERS?

by

LUCY HACKETT

A THESIS

Presented to the Department of Economics
and the Robert D. Clark Honors College
in partial fulfillment of the requirements for the degree of
Bachelor of Arts

June 2015

An Abstract of the Thesis of

Lucy Hackett for the degree of Bachelor of Arts
in the department of Economics to be taken June 2015

Title: ENGLISH PROGRAMS IN OREGON: HELPING OR INHIBITING ENGLISH LANGUAGE LEARNERS?

Approved: ____

Professor Glen Waddell

English Language Learners (ELL's) face challenges to their educational success because of the added difficulty of non-fluency in the English Language. This paper studies the effect that English programs in Oregon public schools have on the educational outcomes of ELL's. A regression discontinuity design was utilized to test the effect that treatment with English programming has on ELL students' outcomes on Math and Reading assessments, and revealed either no economically significant or a negligibly positive effect from English programming on either Reading or Math scores.

Thank you to my family for supporting and loving me when I was unsupportable and unlovable.

Thank you also to Glen Waddell, for challenging me and pushing me forward over the last three years, even though I ended up taking none of his advice.

Table of Contents

Index of Graphs and Figures

I. Introduction

English Language Learners (ELL's) are students whose native language or whose primary language spoken in the home is not English. These students, because of their unfamiliarity with the English language, which can range from not having the ability to speak English to only slight deficits in vocabulary and reading comprehension, face challenges in their education due to their lack of English language proficiency.

Several laws, court cases, and state guidelines mandate that public schools respond to the challenge presented by these students by offering English language assistance to students who need it. The precedent of legislation around language minority students began in the 1960's with Title VI of the Civil Rights Act of 1964. This act prohibits discrimination based on race, color or national origin in programs that receive federal financial assistance (Portland Audit 2010). The Supreme Court interpreted this legislation to mean that school programs are required to meet the lingual needs of individuals whose primary language is not English. The Supreme Court, with *Lau vs. Nichols* (1974), affirmed that language minority students must receive special attention and educational programs to preserve their equal access to education. Other legislation, such as the No Child Left Behind Act of 2001, also continued this precedent by establishing more programs and funding for language minority students. Several districts in Oregon, such as Portland Public School District, have been in violation of various state and federal mandates concerning ELL education such as Title VI, the No Child Left

Behind Act and Oregon State Statutes and Administrative Rules, for several years now (Portland SD Audit 2010).

This research will suggest that ELL students are experiencing a negligible or nonexistent treatment effect from English programs. Because this effect is statistically zero or very small, lowering the threshold to qualify for English programs and allocating resources more intensively to the students who experience the greatest need may have minimal costs from a policy standpoint.

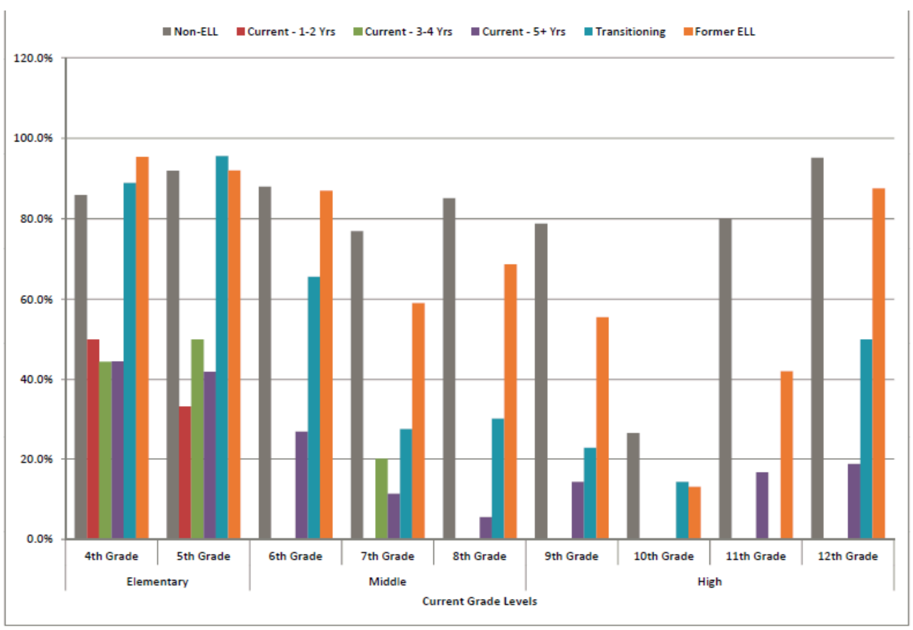
Importance

These questions are important and will grow in importance as the make-up of students in Oregon, and more broadly in the United States, changes to include more language minority students. Data from 2002 to 2012 from the US Department of Education from both Oregon and the nation at large show that the percentage of students whose primary language is not English has grown steadily since 2002 (USDOE, National Center for Education Statistics 2012). The number of students that this affects is large; in 2012, 11.3% of students in Oregon participated in English Language (EL) programs, which amounts to 63,790 students (US DOE 2012). As this number grows, the need for a more effective allocation of resources will increase in importance.

Though 11.3% is clearly a minority of students, it is a relatively large fraction of students to qualify for a specific service. Students with learning disabilities make up only 4.8% of the student body nationally, and about 6.9% of students in Oregon participate in Talented and Gifted programs (USDOE 2012). Additionally, English Language Learners are the fastest growing subpopulation of the student body in the United States (Tracy 2010), and if this trend follows the pattern it has followed over the past 50 years, the need to efficiently allocate the resources that we dedicate to language assistance will certainly not decrease.

ELL students also experience lower educational outcomes than their counterparts who are fluent in English. As can be seen in Figure 1 below, which illustrates reading outcomes from Tigard Tualatin School District in 2013, the gaps in achievement experienced by ELL's worsen with higher grade levels until the 12th grade, where the gap closes. This can be explained by the ELL students that drop out of school, leaving behind only their ELL counterparts who are achieving nearly at the level of English-speaking students.

Figure 1: Percent of Students meeting Reading OAKS Standards, TTSD 2013



The disparities depicted in Figure 1 raise questions about the causes of the gaps present. Especially interesting are the significant differences between non-

ELL students and former ELL students. It is important to note that this figure is not conditioned on any factors and therefore does not capture any differential behavior. Specifically, drop out behavior can explain some of the variation seen such as the sudden rise in the percent of former ELL students meeting OAKS standards in the 12^{th} grade. Though one could speculate about many potential causes, this paper will focus on English programs and the outcomes that English language education is achieving or is failing to achieve.

The programs that English Language Learners are treated with broadly fall into one of three categories. One is Bilingual Education (BE), in which language minority students are separated from native English speakers and receive their instruction partially in their native language and partially in English in order to co-develop English language skills and content learning. The advantage, pedagogically, is generally believed to be that these students acquire English skills without losing valuable instruction time and academic material. This type of program may also benefit language minority students by allowing them to preserve and possibly even enhance the language skills that they possess in their native language.

In Dual Immersion programs, ELL's and native English speakers are instructed in a bilingual setting so that both groups of students become bilingual. These programs are preferred because ELL's are not pulled from class and therefore do not miss instruction time. Additionally, there may be positive externalities to a diverse student body in a setting where the objective for all students is to learn a second language and culture.

Finally, in English as a Second Language (ESL) programs, ELL's are specifically taught English grammar and vocabulary as a second language. This may take the form of ELL students being pulled from class to receive special instruction in another classroom. Or, schools may hire an ESL aid that assists ELL's within the framework of the classroom. Within these categories, the details of the structures may vary, but these summarize the three ways that ELL's receive English assistance.

In Oregon, upon entering the public school system, students who are identified as language minority or potentially ELL are given an English assessment known as the ELPA (English Language Proficiency Assessment). This identification could occur based on an entry survey of the student and their family, or when language barriers arise in a classroom. Based on state and district standards, students are classified into categories of English proficiency based on their ELPA score. If a student's score is above the threshold for "passing", they do not receive English assistance and receive the same educational treatment as native English speakers. If a student does not achieve this threshold on the ELPA, they are classified as an ELL and qualify for English programs. There are exceptions to the process described above, which I will address later in this paper. Though I am not interested in the admissions process itself, I will use this as the basis for my empirical analysis of the effectiveness of English language programs with a regression discontinuity model.

II. Data and Methodology

I utilize data from the Oregon Department of Education (ODE) at the student level. The data come from school districts from all over the state and contain observations for all students in Oregon Public School systems that had taken the ELPA. I am drawing on data from various data collections, including the LEP Collection and the Student Academic Summary Collection. Together, these data provide me with variables such as score on the ELPA, scores on the OAKS Reading and Math exams, language spoken at home, race or ethnicity, economic disadvantage and gender. Because the LEP data collection begins in 2007, I examined recent data from the 2007/2008 school year on.

The sample that was utilized includes students who were not necessarily reported to have taken both the Reading and Math exams. About 4000 observations in the data set did not have both tests recorded. Therefore, the results are not sensitive to the unobserved heterogeneity driving some students to take one test but not the other. However, because 4000 is relatively small compared to the 175,000 plus observations of the Reading and Math samples and the results are not significantly different when these students are excluded, I utilize all test takers whether they have scores for both recorded or not.

This paper will utilize a regression discontinuity design to examine the effect of English program treatment on academic outcomes. There has been resurgence in the use of regression-discontinuity (RD) designs by empirical researchers since the late 1990s. This approach to estimating causal effects is often preferred to all other non-experimental strategies (Cook 2008; Lee and Lemieux 2010) as RD designs usually entail perfect knowledge of the selection process and require comparatively weak assumptions (Hahn, Todd, and van der Klaauw 2001; Lee 2008). Several studies support this view, and have shown that regression discontinuity designs and experimental studies produce similar estimates. RD designs also offer an appealing intuition—so long as characteristics related to outcomes are smooth around the treatment threshold, we can reasonably attribute differences in outcomes across the threshold to the treatment.

In the model used in this research, the assumption that justifies a causal interpretation is that individuals are similar around the ELPA cutoff that assigns treatment to some students and not others. That is, any difference in outcomes between students who scored within a few points of each other on the ELPA but ended up on either side of the treatment-assignment cut-off, and are therefore very different in how they are treated, we attribute to treatment. This is so long as there are not other attributes that also shift at that same threshold. For example, our prior is that there should not be significant differences in English ability or other factors among kindergarten students who score a 506 and other kindergarten students who score a 508. The 508 cohorts therefore offer reasonable control groups against which to measure the effect of treating the 506 cohorts. Because the difference in score is slight, it is reasonable to think that though the students are not truly random in the sense that they were chosen randomly for a randomized controlled experiment, they are practically nearly random in reference to the students who score similarly to them on the ELPA.

The model used to estimate this difference was created using as the dependent variable scores received on Reading and Math assessments, called *read* and *math*. The basic model regressed these variables on the variable *elpa_plusminus*, which recorded a student's score in relation to the cut off ELPA score. For example, if the threshold for a kindergarten student was 508, a score of 507 would take on a value of -1 for that student. In order to compare the outcomes for students who scored above the threshold, and were therefore not treated, and those who failed to achieve the threshold and received treatment, the regressions take the form of Equations 1 and 2 below,

$$read_{it} = \beta_0 + treat_i * \beta_1 + ELPA_{it} * \beta_2 + ELPA_{it}^2 * \beta_3 + ELPA_{it} * treat_i * \beta_4 + ELPA_{it}^2 * treat_i * \beta_5 + u_{it} \quad (1)$$

$$math_{it} = \beta_0 + treat_i * \beta_1 + ELPA_{it} * \beta_2 + ELPA_{it}^2 * \beta_3 + ELPA_{it} * treat_i * \beta_4 + ELPA_{it}^2 * treat_i * \beta_5 + u_{it} \quad (2)$$

where the variable $treat_i$ takes on a value of 0 if the student achieves or surpasses the threshold on the ELPA and a value of 1 if a student scores below the threshold. The regressions were run on the reading and math scores achieved by each student. Therefore, the coefficient on $treat_i$ measures the difference in math scores between students treated with English programming and those not treated.

Though test scores are not inherently what we may be interested in as metrics of academic success, they are a proxy for a student's academic achievement, and as such serve in this research as a litmus test for students' academic outcomes.

Underlying Assumptions

As stated above, the primary assumption in a regression discontinuity model is that students are locally pseudo-random. In order to support this, I first establish that there are no "jumps" in other characteristics of individuals on either side of the cut off score. Any attributes that are not smooth through the treatment threshold could act as a confounding factor, potentially explaining any difference in outcomes as well as the treatment itself would. For example, if students who barely failed the ELPA were significantly less likely to be economically disadvantaged than students who barely passed, economic disadvantage would be a reasonable explanation to explain why outcome measures among ELPA passers were higher. To check for this, I confirm the smoothness of variables for race (Hispanic, White, Black and Asian/Pacific Islander), gender, and economic disadvantage across the threshold. The results of these tests will be discussed in the "Analysis" section later in this paper.

Before proceeding to the formal analysis, note that perfect compliance with treatment is unlikely. That is, it is unlikely that every individual who meets the treatment criterion receives treatment and that every individual that does not is denied treatment. For example, a teacher may recommend that a student who passed the ELPA participate in English assistance if they feel that the student is unable to succeed in the classroom because of their English ability. On the other hand, parents always have the right to refuse services in public school systems, so a parent may opt their child out of an English program even if they classify as ELL.

It is not difficult to imagine a parent that would prefer that their child remain in the classroom with native English speaking students instead of being taught separately with other language minority students. If such parents are also different in their other unobservable attributes that have the potential to drive differences in outcomes (such as educational achievement), we would likely misidentify the true treatment effect.

In considering treatment compliance, previous literature on the subject and the analysis that was possible from the data at hand were utilized. Previous research of ELL students in education has shown that compliance tends to be high for English programs. In California for example, compliance in ELL reclassification has been estimated at a minimum of eighty percent (Robinson 2011). In Tigard Tualatin School District in Oregon, a district with a particularly high number of ELL's (over 12,000 students), they have reported an average program waiver of 6.7% (Tigard Tualatin 2013). That is, 6.7% of students who qualified as ELL's refused additional English services. This rate is reflected in the data that I analyzed, which reports an even lower "Not Participating" rate of 3.48%. Because of the way that the data were inputted and organized, it is not possible to ascertain the rate at which teachers put students who passed the ELPA in the treatment group, overriding their ELPA score.

Because of this consideration, this analysis effectively measures the intent to treat effect for students who score below the threshold. Because it cannot be said for certain that students were treated if they failed to reach the cut-off score, it does not measure the actual treatment effect, rather the effect of the intent to treat created by the policy-established thresholds.

III. Reading and Math Analysis

The overall effect that treatment had on Math and Reading test scores was mixed in magnitude and significance depending on grade level. The point estimates of the means for Reading and Math scores, as can be seen in Table 1 below, at first glance indicate that scores were hurt by treatment for most grade levels.

Table 1: Reading and Math Means by Grade

	Reading Means		Math Means	
	Not Treated	Treated	Not Treated	Treated
Kindergarten	215.7568	205.5	214.8352	203.6215
1st Grade	--	--	--	--
2nd Grade	206.6667	296.5	202.3333	200.5
3rd Grade	215.3454	227.4012	213.2747	205.7982
4th Grade	220.0636	210.2299	221.3235	211.6684
5th Grade	222.5517	212.6633	225.6143	216.2381
6th Grade	226.1871	216.4732	227.1684	217.1362
7th Grade	231.1663	221.2832	232.9957	224.0895
8th Grade	231.8148	222.1206	234.7652	225.5297
9th Grade	234.103	225.8459	232.4493	225.0715
10th Grade	235.2725	226.7128	233.1114	226.2169
11th Grade	237.1648	226.7291	235.0123	226.2275
12th Grade	232.9449	226.5979	230.3413	225.1198

However, the results for Reading outcomes, summarized in Table 2 below, suggest that the lower point estimates for the treated cohort are misleading. The treatment effect varies according to the educational level at which a student is treated with an English program. The only statistically significant results at the 5% level appear for Grades 4 through 7 and 12, and the treated group in each of these cohorts experienced a positive effect on Reading scores from treatment. The statistically significant coefficients represent increases of 0.32% for fourth grade, 0.48% for fifth grade, 0.37% for sixth grade, 0.55% for seventh grade, and 1.87% for twelfth grade. This indicates that these programs are indeed having sparse positive effects on Reading outcomes for ELL's. However, other grade levels experienced a non-statistically significant effect, indicating a lack of any effect at all from treatment with an English program.

Table 2: Reading Outcomes by Grade

VARIABLES	(1) Overall	(2) Grade 3	(3) Grade 4	(4) Grade 5	(5) Grade 6	(6) Grade 7
Treated	0.405	-1.001	0.383**	0.626***	0.463***	0.718***
	(0.322)	(2.403)	(0.151)	(0.154)	(0.166)	(0.187)
ELPA	0.558***	-0.253	0.590***	0.523***	0.440***	0.503***
	(0.0598)	(0.628)	(0.0502)	(0.0463)	(0.0474)	(0.0526)
$ELPA^2$	-0.00454	0.0242	-0.00175	-0.00467	0.00126	-0.00548
	(0.00349)	(0.0289)	(0.00413)	(0.00361)	(0.00348)	(0.00376)
ELPA*Treated	0.411***	1.438**	0.195***	0.249***	0.281***	0.190***
	(0.0821)	(0.668)	(0.0546)	(0.0516)	(0.0520)	(0.0562)
$ELPA^2$*Treated	0.0327***	0.0399	0.00518	0.0105***	0.00485	0.0124***
	(0.00408)	(0.0299)	(0.00423)	(0.00372)	(0.00357)	(0.00381)
Gender	-0.661***	-3.397***	0.691***	0.581***	-0.268***	-0.627***
	(0.200)	(0.901)	(0.0653)	(0.0736)	(0.0807)	(0.0868)
Black	-2.065***	-7.842***	-0.779***	-1.393***	-0.997***	-0.776**
	(0.249)	(1.155)	(0.272)	(0.285)	(0.333)	(0.339)
Hispanic	5.202***	23.40***	-0.117	-0.473***	0.186	-0.212
	(0.169)	(0.703)	(0.113)	(0.121)	(0.139)	(0.144)
Asian/Pac. Is.	0.893***	1.728***	0.964***	0.658***	1.064***	0.808***
	(0.139)	(0.519)	(0.148)	(0.165)	(0.189)	(0.193)
Econ. Disad.	0.297	2.753***	-0.610***	-1.007***	-0.949***	-0.938***
	(0.205)	(0.962)	(0.138)	(0.139)	(0.167)	(0.158)
Constant	223.1***	208.6***	217.4***	219.3***	222.7***	227.6***
	(0.316)	(2.107)	(0.188)	(0.189)	(0.222)	(0.231)
Observations	175,433	38,178	33,807	26,589	20,045	16,874
Mean	224.33	222.80	215.39	218.96	223.15	228.37

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2 Continued

VARIABLES	(7) Grade 8	(8) Grade 9	(9) Grade 10	(10) Grade 11	(11) Grade 12
Treated	0.240	0.166	0.451	0.244	2.353***
	(0.201)	(0.448)	(0.281)	(0.389)	(0.576)
ELPA	0.308***	0.663***	0.421***	0.169	0.898***
	(0.0498)	(0.187)	(0.110)	(0.130)	(0.193)
$ELPA^2$	0.000748	-0.0322**	0.0119	0.00804	-0.0232**
	(0.00338)	(0.0161)	(0.00911)	(0.00964)	(0.0114)
ELPA*Treated	0.334***	-0.0549	0.273**	0.610***	-0.187
	(0.0553)	(0.189)	(0.112)	(0.134)	(0.198)
$ELPA^2$ *Treated	0.00451	0.0387**	-0.00360	0.00157	0.0315***
	(0.00348)	(0.0161)	(0.00912)	(0.00968)	(0.0115)
Gender	-0.442***	-0.0879	-0.105	-0.446***	-0.0663
	(0.0923)	(0.145)	(0.102)	(0.160)	(0.215)
Black	-0.979***	-1.282**	-0.986***	-0.577	-1.181**
	(0.337)	(0.533)	(0.298)	(0.407)	(0.540)
Hispanic	-0.0808	-0.333	0.233	-0.273	-0.688*
	(0.160)	(0.235)	(0.172)	(0.282)	(0.357)
Asian/Pac. Is.	0.636***	0.187	0.741***	0.993***	-0.185
	(0.212)	(0.314)	(0.214)	(0.312)	(0.400)
Econ. Disad.	-1.290***	-0.431*	-0.511***	0.304	-0.0460
	(0.180)	(0.250)	(0.163)	(0.230)	(0.302)
Constant	229.3***	232.5***	233.6***	234.2***	232.3***
	(0.277)	(0.493)	(0.316)	(0.445)	(0.607)
Observations	14,778	5,639	11,623	5,135	2,761
Mean	229.14	231.65	232.06	234.82	230.26

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Results from the "Overall" column control for grade level; coefficients for grade not listed above. The variable *ELPA* represents the difference between a student's score and the proficiency threshold. (ELPA score – ELPA cut-off).

In Math, the only significant coefficients were for the overall effect on students of all grades with a coefficient of 0.232 points and for the third grade (0.612 points), significant at the 1% and 5% levels, as can be seen in Table 3 below. The magnitude of this result is small; it represents a 0.104% and 0.29% increase in Math scores from untreated to treated. This indicates that in most grade levels ELL students fare the same on Math examinations regardless of treatment with an English program. And where they do improve, the treatment effect is very slight. So, the effect on Math scores is even smaller than that seen for Reading scores, and the effect is less persistent than the effect for Reading. This is intuitive given the language aspect of English acquisition.

Table 3: Math Outcomes by Grade

VARIABLES	(1) Overall	(2) Grade 3	(3) Grade 4	(4) Grade 5	(5) Grade 6	(6) Grade 7
Treated	0.232***	0.612**	0.185	0.171	0.380*	0.0966
	(0.0854)	(0.257)	(0.190)	(0.194)	(0.226)	(0.230)
ELPA	0.524***	0.388***	0.620***	0.499***	0.541***	0.320***
	(0.0276)	(0.0857)	(0.0526)	(0.0660)	(0.0720)	(0.0686)
$ELPA^2$	-0.00219	-0.00740	-0.000908	-0.000488	-0.00146	0.00664
	(0.00210)	(0.00549)	(0.00374)	(0.00535)	(0.00556)	(0.00515)
ELPA*Treated	0.0764***	0.130	0.119**	0.106	0.143*	0.234***
	(0.0289)	(0.0880)	(0.0573)	(0.0706)	(0.0759)	(0.0729)
$ELPA^2$*Treated	0.00723***	0.00700	0.00694*	0.00482	0.0102*	-2.23e-05
	(0.00212)	(0.00553)	(0.00384)	(0.00542)	(0.00562)	(0.00520)
Gender	1.462***	1.905***	2.082***	1.602***	1.238***	0.750***
	(0.0367)	(0.0808)	(0.0853)	(0.0922)	(0.103)	(0.109)
Black	-2.879***	-1.983***	-2.874***	-2.901***	-3.314***	-3.104***
	(0.142)	(0.356)	(0.364)	(0.354)	(0.433)	(0.392)
Hispanic	-1.041***	-0.597***	-1.034***	-0.902***	-0.560***	-1.265***
	(0.0614)	(0.133)	(0.147)	(0.154)	(0.176)	(0.181)
Asian/Pac. Is.	3.517***	2.231***	2.099***	2.906***	4.075***	4.251***
	(0.0901)	(0.185)	(0.206)	(0.230)	(0.272)	(0.284)
Econ. Disad.	-1.332***	-0.984***	-1.185***	-1.617***	-2.337***	-1.989***
	(0.0720)	(0.151)	(0.184)	(0.188)	(0.222)	(0.224)
Constant	231.7***	213.2***	218.9***	222.4***	223.8***	230.7***
	(0.122)	(0.272)	(0.240)	(0.249)	(0.301)	(0.307)
Observations	179,006	38,549	34,309	27,059	20,469	17,350
Mean	223.19	208.63	216.69	222.14	223.96	230.41

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3 Continued

VARIABLES	(7) Grade 8	(8) Grade 9	(9) Grade 10	(10) Grade 11	(11) Grade 12
Treated	0.0559	0.0605	0.129	-0.536	-0.735
	(0.280)	(0.602)	(0.369)	(0.512)	(0.913)
ELPA	0.549***	0.493**	0.434***	0.283	0.349
	(0.0760)	(0.236)	(0.116)	(0.181)	(0.325)
$ELPA^2$	-0.00724	0.0113	0.0169**	0.00489	-0.00146
	(0.00517)	(0.0192)	(0.00792)	(0.0131)	(0.0266)
ELPA*Treated	-0.0827	0.0479	0.118	0.277	0.0468
	(0.0812)	(0.239)	(0.119)	(0.185)	(0.333)
$ELPA^2$*Treated	0.0110**	-0.00474	-0.0104	0.00148	0.00324
	(0.00525)	(0.0192)	(0.00796)	(0.0131)	(0.0267)
Gender	1.039***	1.210***	1.068***	0.676***	0.148
	(0.127)	(0.188)	(0.141)	(0.201)	(0.348)
Black	-3.519***	-2.579***	-3.247***	-4.113***	-6.060***
	(0.453)	(0.736)	(0.449)	(0.630)	(0.858)
Hispanic	-1.726***	-0.263	-1.609***	-2.197***	-2.712***
	(0.212)	(0.313)	(0.240)	(0.340)	(0.581)
Asian/Pac. Is.	4.836***	5.932***	4.716***	4.603***	2.443***
	(0.335)	(0.486)	(0.340)	(0.437)	(0.775)
Econ. Disad.	-1.880***	-1.321***	-0.931***	0.0477	-1.197**
	(0.260)	(0.340)	(0.235)	(0.316)	(0.499)
Constant	232.0***	230.6***	232.8***	233.2***	233.7***
	(0.374)	(0.674)	(0.418)	(0.580)	(1.002)
Observations	15,163	6,139	11,736	6,249	1,979
Mean	232.15	230.34	230.49	232.78	228.61

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Results from the "Overall" column control for grade level; coefficients for grade not listed above. The variable *ELPA* represents the difference between a student's score and the proficiency threshold. (ELPA score – ELPA cut-off).

When examining this effect through the mean of Math and Reading scores overall, there is no statistically significant effect on either type of test score produced by the treatment (Table 4). Additionally, upon examining the mean for each grade level individually, treatment appears to have no effect on Reading or Math scores for any grade (Figures 5, 6), further indicating that there is no significant treatment effect statistically or in magnitude on Math or Reading examinations.

Table 4: Overall Reading and Math Mean Effect

VARIABLES	(1) Math	(2) Reading
Treated	0.0765	-1.694*
	(1.036)	(0.990)
ELPA	0.0723	0.0614
	(0.195)	(0.114)
$ELPA^2$	0.00524	0.00452
	(0.00634)	(0.00371)
ELPA*Treated	0.914***	0.387**
	(0.200)	(0.191)
$ELPA^2$*Treated	0.0105	0.0130**
	(0.00642)	(0.00582)
Constant	224.1***	223.4***
	(0.978)	(0.550)
Observations	79	79
R-squared	0.870	0.610

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Discussion of Continuity

Regressions measuring the “jumps” in covariates utilized Equation 3 (the variable $covariate_i$ represents the mean of a dummy indicator for a covariate analyzed for each ELPA score as an example):

$$covariate_i = \beta_0 + treat_i * \beta_1 + ELPA_{it} * \beta_2 + ELPA_{it}^2 * \beta_3 + ELPA_{it} * treat_i * \beta_4 + ELPA_{it}^2 * treat_i + u_{it} \quad (3)$$

All returned statistically insignificant results at the 5% level, except for Black students, who are somewhat more represented on the treatment side of the threshold (Table 5). However, the coefficient for Black students is small (0.00827), making it doubtful that the treatment effect can be explained by this discontinuity. The results of the tests for continuity through the threshold of other factors appear in Table 5 below.

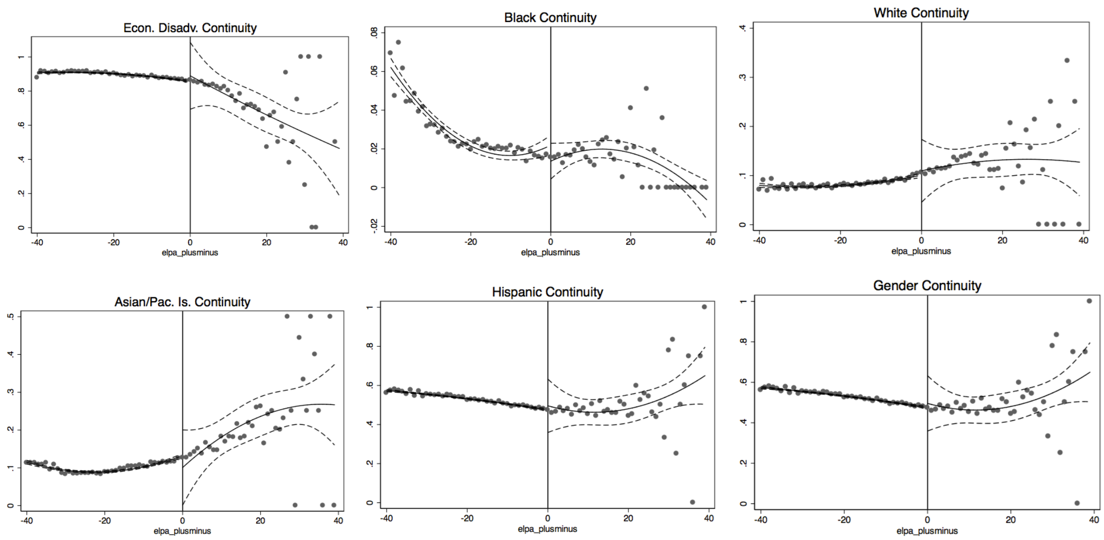
Table 5: Continuity of Covariates

VARIABLES	(1) Hispanic	(2) Gender	(3) Asian/Pac. Islander	(4) White	(5) Black	(6) Economic Disad.
Treated	-0.0358	-0.0209	0.0316	-0.00908	0.00827**	-0.0333
	(0.0390)	(0.0355)	(0.0290)	(0.0156)	(0.00333)	(0.0311)
ELPA	-0.0130	-0.00573	0.00940	0.00180	0.000975*	-0.0120
	(0.00826)	(0.00783)	(0.00612)	(0.00344)	(0.000513)	(0.00780)
$ELPA^2$	0.000241	0.000248	-0.000132	-3.45e-05	-3.81e-05***	2.88e-05
	(0.000265)	(0.000269)	(0.000198)	(0.000115)	(1.24e-05)	(0.000309)
ELPA*Treated	0.00680	0.00248	-0.00553	-0.000190	7.22e-05	0.00871
	(0.00827)	(0.00784)	(0.00612)	(0.00346)	(0.000605)	(0.00782)
$ELPA^2$*Treated	-0.000388	-0.000268	0.000218	6.18e-05	8.94e-05***	-8.00e-05
	(0.000265)	(0.000269)	(0.000198)	(0.000116)	(1.55e-05)	(0.000309)
Constant	0.759***	0.496***	0.101***	0.110***	0.0136***	0.890***
	(0.0388)	(0.0355)	(0.0289)	(0.0154)	(0.00248)	(0.0309)
Observations	79	79	79	79	79	76
R-squared	0.424	0.147	0.447	0.183	0.739	0.459

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

There is no statistical discontinuity in the ratio of Hispanic, Asian/Pacific Islander, White, Male/Female, or Economic Disadvantage (Table 5). Figure 2 below illustrates graphically the smoothness of the covariates studied.

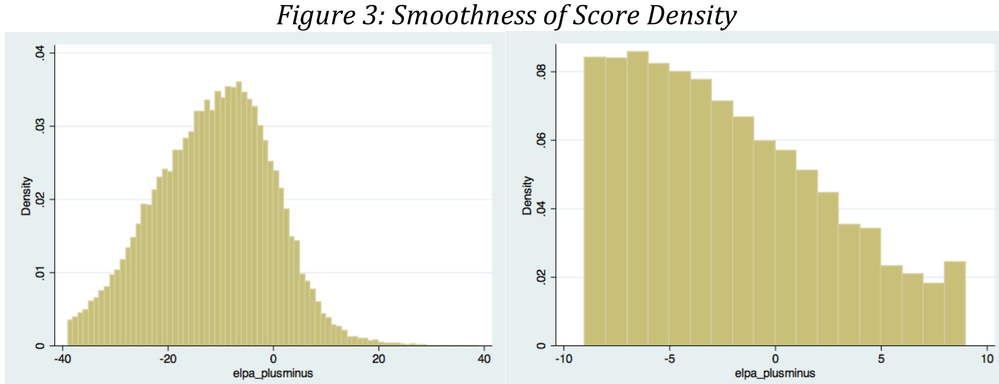
Figure 2: Smoothness of Covariates



As should be the case if the regression discontinuity design is to identify the causal effect of treatment, the lines appear continuous in spite of the fact that the ratio of each cohort changes along the distribution of ELPA scores.

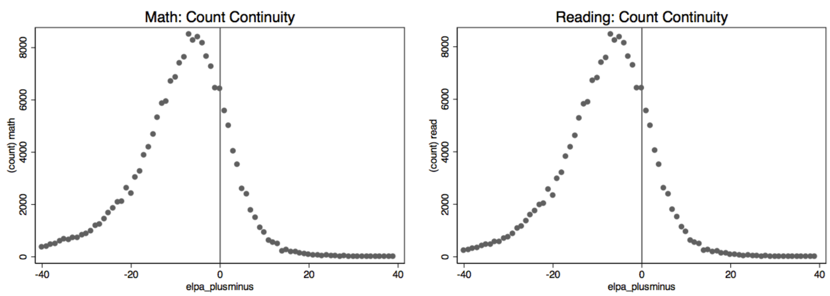
Smoothness was also examined for the density of observations of ELPA scores. As can be seen below, it appears that there are no discontinuities around the threshold for ELPA scores, which would suggest that the policy was being gamed and/or low compliance.

Figure 3: Smoothness of Score Density



Upon closer examination of continuity for the density of scores for Math and Reading, the densities also appear to be smooth. Below, the densities of Math and Reading scores are plotted by ELPA score. There is no stacking on one side of the threshold, and the two sides of the distribution appear fairly symmetrical, indicating that students are not somehow scoring disproportionately on one side of the threshold, which would undermine the treatment effect seen.

Figure 4: Continuity of Density



IV. Further Heterogeneity Analysis

Examining the data further, one finds disparities in the rates at which different populations of students are affected by treatment with English programs. We have examined heterogeneity across grade level; now this research will examine the heterogeneity experienced by different cohorts of students. A cursory examination of the point estimates for the means of different cohorts above and below the threshold indicate that the test scores of each group tended to be lower given treatment, as illustrated in Table 6.

Table 6: Reading and Math Means by Cohort

	Reading Means		Math Means	
	Not Treated	Treated	Not Treated	Treated
Hispanic	221.4265	226.2988	226.3246	215.2616
Black	224.8041	213.8453	224.0899	212.9886
White	228.2107	214.6734	229.5604	216.927
Asian/Pacific Is.	229.9972	216.2411	234.2431	220.2774
Economic Disad.	226.5565	220.2705	227.0295	215.5611
Gender	226.8111	219.455	228.3611	216.505

However, as outlined in Tables 7.1 to 7.5, different ethnic groups experienced differing effects from treatment with English assistance. White and Asian and Pacific Islander ELL students were more positively affected than other ELL students in Reading and were hardly affected (if affected at all) in their Math scores, while Hispanic ELL students scored disproportionately high on Math assessments but scored lower on Reading than their counterparts when treated.

Regressions analyzing the effect on students who have a 1 for the dummy variable *Economic Disadvantage* showed similar heterogeneity, as can be seen in Table 7.5 below. Students with economic disadvantages scored lower in Reading and higher on Math assessments when treated than their counterparts not marked as economically disadvantaged.

Table 7.5: Economic Disadvantage Heterogeneity

VARIABLES	(1) Math: No Econ. Dis	(2) Math: Econ. Dis.	(3) Reading: No Econ. Dis.	(4) Reading: Econ. Dis.
Treated	0.709*	1.387***	1.241**	-0.0266
	(0.381)	(0.115)	(0.536)	(0.363)
ELPA	0.888***	0.579***	0.636***	0.310***
	(0.0958)	(0.0366)	(0.0791)	(0.0696)
$ELPA^2$	-0.00896	-0.00865***	-0.00601	0.00243
	(0.00594)	(0.00281)	(0.00512)	(0.00396)
ELPA*Treated	-0.184*	0.354***	0.145	0.311***
	(0.104)	(0.0384)	(0.135)	(0.0926)
$ELPA^2$*Treated	0.0178***	0.0225***	0.0191***	0.0213***
	(0.00608)	(0.00284)	(0.00666)	(0.00456)
Constant	224.8***	222.4***	222.9***	222.7***
	(0.289)	(0.0897)	(0.258)	(0.227)
Observations	18,535	160,471	17,466	157,967
R-squared	0.160	0.202	0.037	0.003

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The possibility for covariation is high here, so this research will not speculate as to the causes of these heterogeneous treatment effects, but further analysis would be necessary to determine the causes of these different outcomes. It is likely that different data would be necessary to conduct that analysis. However, this research suggests that if the goal of English programming is to affect language minority students' academic achievement through assistance with English language skills, it is not being achieved. Furthermore, because different types of students are affected differently by treatment, perhaps different approaches for different types of students would be more effective than a one-size-fits-all treatment.

V. Conclusion

More research is needed to identify the specific causes and dynamics at play with English language programs, but this analysis suggests that programs designed to help ELL students in Oregon are having either a neutral or almost negligible slightly positive effect on the Reading and Math scores of ELL students. In fact, this exercise has precisely identified very small positive effects. English language programs do have a differential effect for students of different ethnicities and income backgrounds, suggesting that something about these groups of students causes them to have distinct outcomes from English programs. Yet, again, these differentials, having been precisely estimated, are economically insignificant or negligible. Overall, this suggests that around the ELPA treatment cut-offs (where regression discontinuity models are well suited to identify the effect of such programs), programs for ELL students are not making a substantial impact in their objective of assisting language minority students academically.

To be clear, this does not support an abandoning of ESL programs. For example, there may be large gains among those with greatest English deficiency, well below the ELPA margin I am able to consider here. What it does support, however, is a more-intensive (i.e., budget neutral) targeting of available resources toward those in greatest need, as I identify negligible costs associated with lowering the ELPA thresholds. To the extent benefits accrue in other than math and reading scores, however, one would worry about lowering ELPA thresholds for treatment. As such, research into other such outcomes would be a worthwhile undertaking.

Appendix A: Regression Tables and Figures by Grade Level

Figure 5: Reading Mean Outcomes by Grade Level

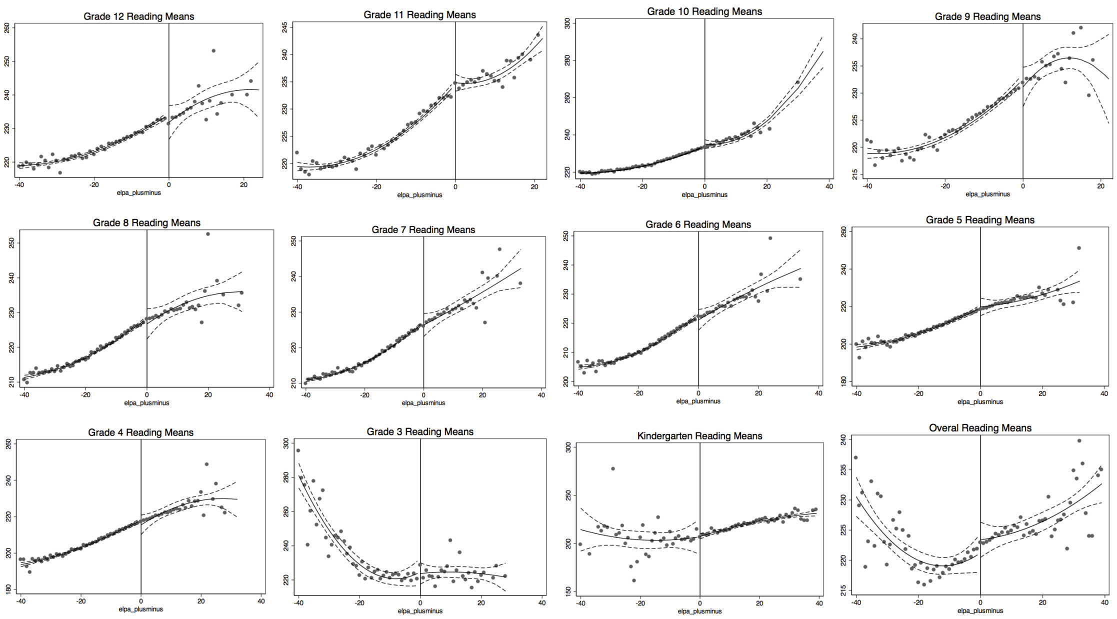
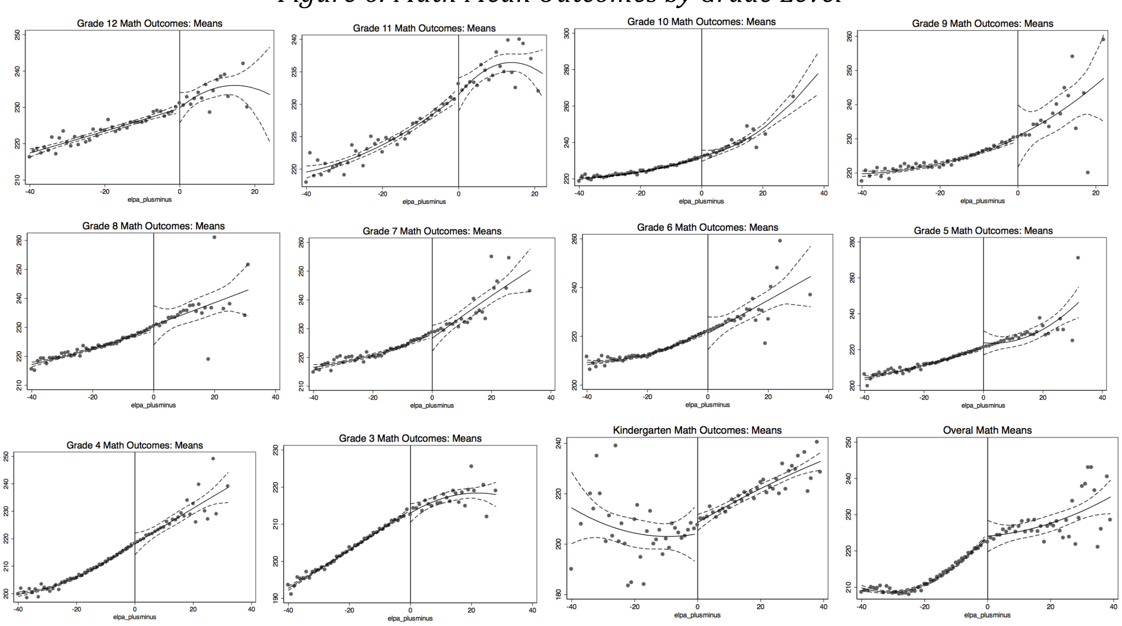


Figure 6: Math Mean Outcomes by Grade Level



Appendix B: Heterogeneity Regressions

Table 7.1: Black Student Heterogeneity Analysis

VARIABLES	(1) Not Black Reading	(2) Black Reading	(3) Not Black Math	(4) Black Math
Treated	0.187	1.505*	1.294***	0.901
	(0.330)	(0.773)	(0.113)	(0.829)
ELPA	0.297***	0.570***	0.570***	0.627***
	(0.0511)	(0.180)	(0.0320)	(0.201)
ELPA^2	-0.00479**	-0.0199*	-0.0159***	-0.0241**
	(0.00240)	(0.0102)	(0.00207)	(0.0107)
ELPA*Treated	0.376***	0.188	0.357***	0.0575
	(0.0782)	(0.196)	(0.0341)	(0.211)
ELPA^2*Treated	0.0303***	0.0283***	0.0299***	0.0303***
	(0.00328)	(0.0104)	(0.00211)	(0.0108)
Constant	222.7***	221.4***	222.6***	221.6***
	(0.192)	(0.581)	(0.0873)	(0.668)
Observations	173,746	4,028	177,005	4,192
R-squared	0.004	0.175	0.184	0.210

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 7.2: Hispanic Student Heterogeneity Analysis

VARIABLES	(1) Not Hispanic Reading	(2) Hispanic Reading	(3) Not Hispanic Math	(4) Hispanic Math
Treated	1.397***	0.0231	0.570**	1.500***
	(0.221)	(0.422)	(0.253)	(0.121)
ELPA	0.584***	0.192***	0.764***	0.470***
	(0.0507)	(0.0667)	(0.0657)	(0.0343)
ELPA^2	-0.0132***	-0.00247	-0.0187***	-0.0160***
	(0.00316)	(0.00304)	(0.00407)	(0.00220)
ELPA*Treated	0.299***	0.482***	-0.0387	0.500***
	(0.0578)	(0.101)	(0.0704)	(0.0366)
ELPA^2*Treated	0.0239***	0.0346***	0.0263***	0.0310***
	(0.00327)	(0.00420)	(0.00413)	(0.00224)
Constant	222.2***	222.8***	224.8***	222.0***
	(0.148)	(0.248)	(0.195)	(0.0933)
Observations	42,092	135,682	43,787	137,410
R-squared	0.178	0.005	0.175	0.196

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 7.3: Asian/Pacific Islander Student Heterogeneity Analysis

VARIABLES	(5) Not Asian/Pac. Islander Reading	(6) Asian/Pac. Islander Reading	(7) Not Asian/Pac. Islander Math	(8) Asian/Pac. Islander Math
Treated	0.0275	1.438***	1.397***	0.445
	(0.360)	(0.293)	(0.112)	(0.404)
ELPA	0.250***	0.631***	0.509***	0.896***
	(0.0569)	(0.0745)	(0.0312)	(0.0993)
$ELPA^2$	-0.00391	-0.0147***	-0.0164***	-0.0211***
	(0.00264)	(0.00457)	(0.00198)	(0.00593)
ELPA*Treated	0.399***	0.191**	0.439***	-0.318***
	(0.0854)	(0.0811)	(0.0334)	(0.107)
$ELPA^2$*Treated	0.0302***	0.0237***	0.0305***	0.0252***
	(0.00357)	(0.00468)	(0.00202)	(0.00604)
Constant	222.7***	222.7***	222.1***	226.5***
	(0.212)	(0.218)	(0.0866)	(0.310)
Observations	159,041	18,733	161,376	19,821
R-squared	0.003	0.257	0.199	0.148

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 7.4: White Student Heterogeneity Analysis

VARIABLES	(1) Not White Reading	(2) White Reading	(3) Not White Math	(4) White Math
Treated	-0.0739	1.453***	1.303***	0.797**
	(0.352)	(0.327)	(0.118)	(0.356)
ELPA	0.275***	0.515***	0.559***	0.616***
	(0.0558)	(0.0762)	(0.0342)	(0.0864)
$ELPA^2$	-0.00437*	-0.0108**	-0.0159***	-0.0164***
	(0.00259)	(0.00475)	(0.00223)	(0.00521)
ELPA*Treated	0.326***	0.482***	0.362***	0.226**
	(0.0829)	(0.0888)	(0.0362)	(0.0955)
$ELPA^2$*Treated	0.0282***	0.0233***	0.0296***	0.0270***
	(0.00345)	(0.00499)	(0.00227)	(0.00538)
Constant	222.8***	221.9***	222.5***	223.8***
	(0.209)	(0.231)	(0.0918)	(0.265)
Observations	162,044	15,730	165,063	16,134
R-squared	0.003	0.232	0.183	0.204

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Bibliography

Bond, T and K. Lang. (2013). "The Black-White Education test-score gap in grades K-7", *NBER*, 7.

Callahan, R. (2005). "Tracking and High School English Learners: Limiting opportunity to learn," *American Educational Research Journal,* 42(2), 305-328.

Chin, Daysal and Imberman (2013). "Impact of Bilingual Education Programs on Limited English Proficient Students and their Peers: Regression Discontinuity Evidence from Texas," *Journal of Public Economics*, 8.

Cook. T.D. (2008). "Waiting for Life to Arrive': A history of the regression-discontinuity design in Psychology, Statistics and Economics," *Journal of Econometrics,* 142(2), 636-654.

"Digest of Education Statistics". National Center for Education Statistics, 2013. *http://nces.ed.gov/programs/digest/.*

"English Learner (EL) Program Plan". Tigard-Tualatin School District, 2013. *http://www.ttsdschools.org/pages/ttsd*.

Fryer, R. and S. Levitt. (2004). "Falling Behind: New evidence on the Black-White achievement gap," *Education Next*, 4(4).

Hahn, J., P. Todd ad W. Van der Klaauw (2001). "Identification and estimation of treatment effects with a regression-discontinuity design," *Econometrica,* 69(1), 201-2019.

Lee, D.S., and Lemieux (2010). "Regression Discontinuity Designs in Economics," *Journal of Economic Literature*.

Robinson, J. (2011) "Using Binding-Score Regression Discontinuity Design with Instrumental Variables Evaluating Criteria for English Learner Reclassification: A Causal-Effects Approach," *Educational Evaluation and Policy Analysis*, 33, 267-284.

Tracy, R. (2010). *Performance Audit of Portland Public Schools.* Portland Public School District.