EDLD 650: Data Analysis and Replication Exercise (DARE) 4

David D. Liebowitz

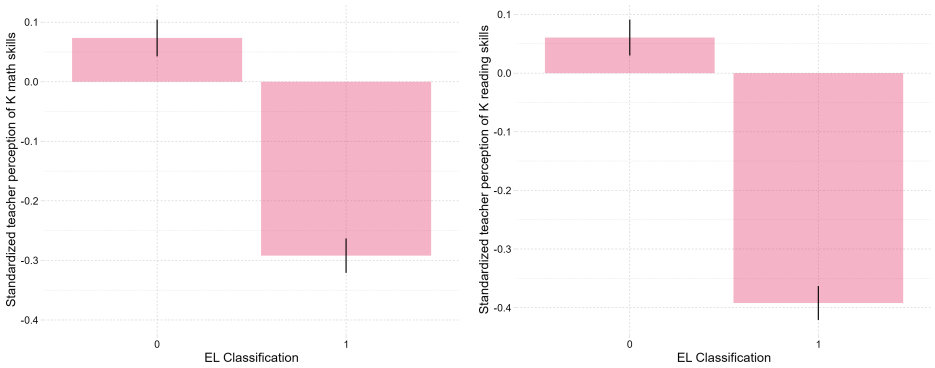
Due: Mar. 3, 2024

A. Baseline differences (3 points)

A1. In the full ECLS-K data, teachers perceive kindergarten students classified as language learners (ELs) as having substantially lower math and language skills than their peers who also predominantly speak a language other than English at home, but are not classified as ELs. In Figure 1, we present visual evidence of the gap in perceived skills. Numerically, teachers' perceptions of EL-classified students (on a mean zero, standard deviation one scale) is 0.365 *SD* units lower for math and 0.453 *SD* units lower in language.

These differences should not be interpreted as the causal effect of being classified as EL on teachers perceptions of students' abilities. These differences could be driven by (a) true (potentially observable) ability differences in students who are classified as EL and those who are not; and (b) unobservable characteristics that cause some students to be classified as EL and others not that also drive teacher perceptions.

A2. In fact, as we document in Table 1, students who are and who are not classified as EL are different across many observable dimensions, even though all students in the sample speak a language other than English at home. Specifically, students classified as EL have lower scores on both English language and content tests, are from lower-income families, are less likely to be from rural areas and are more likely to be Hispanic/Latinx than their non-EL-classified peers. The discrepancies in characteristics of individuals in our treatment and control conditions motivate the matching strategies we employ in the subsequent section.



(a) Math (b) Language

Figure 1: Standardized teacher perception of math and language skills, by EL program assignment

	Not EL (N=945)	EL (N=1221)	Total (N=2166)	p value
Language Score				
				< 0.001
Mean (SD)	17.62 (3.32)	14.55 (4.89)	15.89 (4.54)	
Reading Score				
				< 0.001
Mean (SD)	13.41 (4.57)	10.79 (4.72)	11.94 (4.83)	
K Literacy Score				
				< 0.001
Mean (SD)	-0.49 (0.96)	-1.03 (0.80)	-0.80 (0.91)	
K Math Score				
				< 0.001
Mean (SD)	-0.50 (1.00)	-1.07 (1.00)	-0.82 (1.04)	
SES				
				< 0.001
Mean (SD)	-0.21 (0.91)	-0.69 (0.68)	-0.48 (0.82)	
Female				0.396
Mean (SD)	0.51 (0.50)	0.49 (0.50)	0.50 (0.50)	
Rural school				
				< 0.001
Mean (SD)	0.14 (0.34)	0.08 (0.27)	0.11 (0.31)	
Hispanic/Latinx				
				< 0.001
Mean (SD)	0.50 (0.50)	0.75 (0.44)	0.64 (0.48)	

Table 1: Descriptive statistics by assigned EL status

B. Replication and Extension (7 points)

B1. We begin with a baseline assessment of the probability that a student will be classified as EL, given her observable characteristics. Specifically, we fit the following logistic regression model:

$$ELPRGRM_i = \frac{1}{1 + \exp^{\beta_1 PreLAS_i + \beta_2 EBRS_i + \mathbf{X}_i\theta_i + \epsilon_i}} \quad (1)$$

where we estimate the probability that a student will be assigned to an EL program ($ELPRGRM$) as a non-linear function of their language assessment scores ($PreLAS$) and ($EBRS$) which serve as proxies for the district-administered language test used to determine placement. These are our two primary matching variables. We also include a vector of student demographic characteristics ($\mathbf{X}_i$) that includes their socio-economic status, whether they attend a rural school or not, their gender and their ethnicity (Hispanic/Latinx or not).

In Figure 2, we plot smoothed densities of students' probability of being classified as EL based on their observable characteristics, based on whether they, in fact, received EL services. While there are broad areas of overlap (often referred to as "regions of common support"), there are also substantial regions of both high- and low-probability of EL classification where there are large imbalances in the proportion of students by language classification. This further justifies our matching approach.

B2. We implement a Coarsened Exact Matching (CEM) approach in an effort to identify the causal effect of EL-classification on teachers' perceptions of students' abilities. Our primary matching variables emerge from the legal guidelines of the Elementary and Secondary Education Act that dictate that school districts must use students' home language and English proficiency levels to determine their designation as ELs (or not). Specifically, in a sample of students who speak a language other than English at home, we match on a coarsened value of students' Preschool Language Assessment Scale (Pre-LAS) and English Basic Reading Skill (EBRS) assessments. Additionally, we rely on a coarsened version of students' socio-economic status. For each of these continuous variables, we divide students' scores into quintiles in our sample and match students within quintiles. We also require students to match exactly on gender, attendance in a rural school or not, and Hispanic/Latinx ethnicity (or not). This CEM matching algorithm produces 1627 matched kindergartners, of whom 887 are classified as EL and 740 are not.

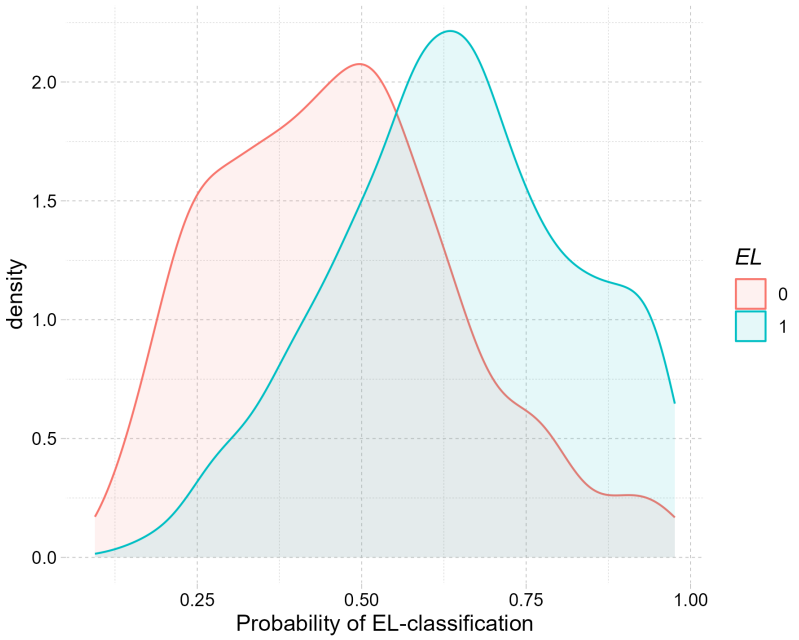


Figure 2: Full sample probability of EL program assignment, by actual EL program assignment

Before proceeding to our estimates of the effects of EL classification, we pause briefly to consider what assumptions must hold in order to interpret our results as supporting causal interpretation. Our central assumption is that, conditional on students' assessed language proficiency, the variation we observe in assignment to treatment (EL) or control (not-EL) conditions depends only on variation in local policy regarding thresholds for EL assignment. Put differently, we assume that students with the same baseline language proficiency and a home language other than English are equal in expectation in their teachers' assessments of their skills, and that no unobservable factors drive categorization of students into either category.

B3. Our CEM matching strategy substantially reduces the observable differences in the characteristics of EL- and non-EL-classified kindergartners and increases the shared region of common support, but it does not produce perfectly balanced treatment and control samples. In Table 2, we present descriptive evidence on the 1627-observation matched sample. We achieve perfect matches on all dichotomous variables (*RURAL*, *FEMALE* and *HISPANIC/LATINX*), and we have removed all statistical differences in students' socio-economic status. We have also eliminated all substantive and significant differences in students' English proficiency scores. Figure 3 provides graphical evidence of the

	Not EL (N=740)	EL (N=887)	Total (N=1627)	adj.pvalue
Language Score				0.786
Mean (SD)	14.97 (5.02)	14.91 (4.98)	14.94 (5.00)	
Reading Score				0.909
Mean (SD)	10.91 (4.89)	10.89 (4.85)	10.90 (4.87)	
SES				0.752
Mean (SD)	-0.66 (0.74)	-0.67 (0.72)	-0.67 (0.73)	
Rural school				1.000
Mean (SD)	0.03 (0.17)	0.03 (0.17)	0.03 (0.17)	
Female				1.000
Mean (SD)	0.47 (0.50)	0.47 (0.50)	0.47 (0.50)	
Hispanic/Latinx				1.000
Mean (SD)	0.77 (0.42)	0.77 (0.42)	0.77 (0.42)	

Table 2: CEM-matched descriptive statistics by assigned EL status

quality of our matches. As compared to Figure 2, we see evidence of much greater regions of common support, but remaining differences in the distributions of EL and non-EL students' probability of receiving the treatment.

While this matching strategy has served to remove much of the bias introduced by baseline characteristic differences between the two groups, it is not perfect and may merit further reconsideration. Our CEM strategy has common support throughout the distribution, so we are still able to generalize to the same population of students who predominantly speak a language other than English at home. However, our common support is not perfect, and we have reduced our sample substantially. We would likely want to vary the width of the bins on which to match our continuous variables or to include more matching variables that better accounted for differences in students categorized as EL and not. The first strategy is preferable as it preserves the criteria-based assignment we have theorized is driving selection into treatment. This allows us to better defend our identification assumptions. However, we risk losing common support within our strata. The second approach will preserve more of our sample, but we stray further from our hypothesized selection process which threatens the assumptions of our overall approach.

B4. As a robustness check, we estimate an alternate Propensity Score

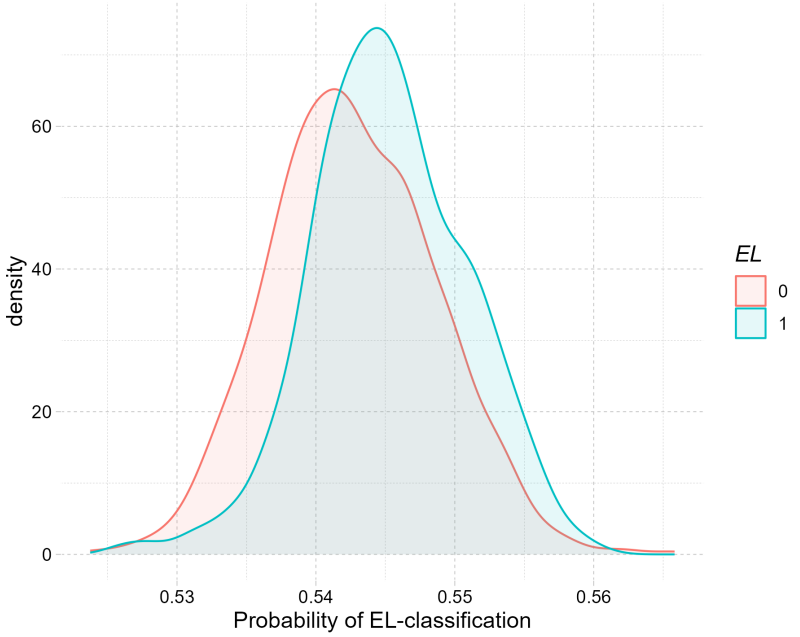


Figure 3: CEM-matched sample probability of EL program assignment, by actual EL program assignment

Matching (PSM) algorithm in which we match treated (EL) students to their nearest-neighbor by propensity score (with replacement) to construct a separate matched sample. We present analogous evidence in Table 3 on the quality of our matches and Figure 4 on the region of common support. Our PSM approach produces roughly equal quality matches, with around the same sample size. However, while we fail to reject the null hypothesis, some of our balancing variables are substantively different across students who are and are not classified as EL (in particular their family-income levels, which differ by 7 percentage points). In all cases, we found a match for each of our treated units, and we did so without requiring the sample to become imbalanced because none of the treated units fell outside the region of common support.[1]

B5. In Table 4, we present results from OLS, CEM and PSM estimates of the effects of EL-classification on teacher perceptions of student abilities in math and reading in kindergarten. Models 1 and 5 estimate the simple bivariate relationship and are equivalent to the mean differences we present in A1. In Models 2 and 6, we introduce a vector of

[1]Additionally, we did not specify a caliper-width outside of which we would refuse to allow students to match which would have led us to drop some of the treated units from our sample.

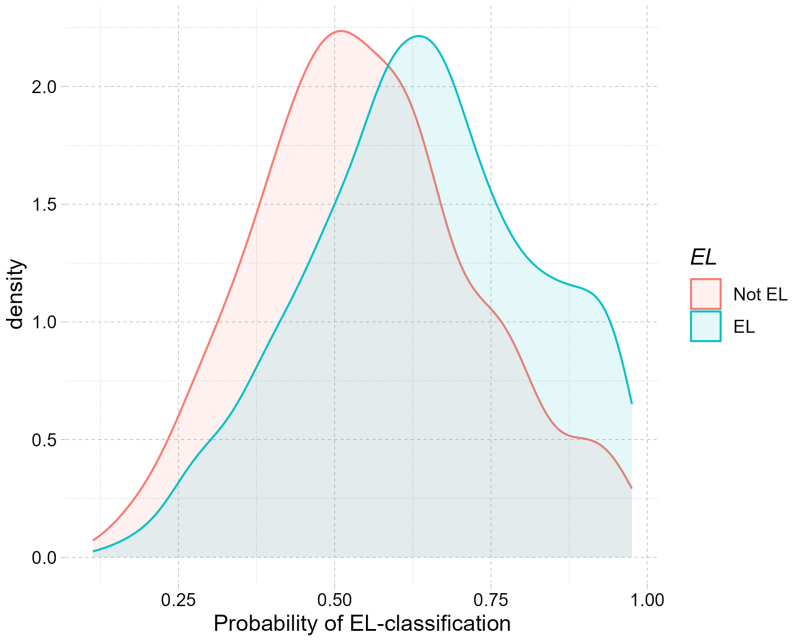


Figure 4: PSM-matched sample probability of EL program assignment, by actual EL program assignment

	Not EL (N=480)	EL (N=1221)	Total (N=1701)	adj.pvalue
Propensity score				0.994
Mean (SD)	0.64 (0.18)	0.64 (0.18)	0.64 (0.18)	
Language Score				0.488
Mean (SD)	14.90 (4.98)	14.55 (4.89)	14.65 (4.92)	
Reading Score				0.261
Mean (SD)	10.79 (4.79)	10.79 (4.72)	10.79 (4.74)	
SES				0.560
Mean (SD)	-0.76 (0.66)	-0.69 (0.68)	-0.71 (0.68)	
Rural school				0.153
Mean (SD)	0.06 (0.23)	0.08 (0.27)	0.07 (0.26)	
Female				0.608
Mean (SD)	0.47 (0.50)	0.49 (0.50)	0.48 (0.50)	
Hispanic/Latinx				0.890
Mean (SD)	0.77 (0.42)	0.75 (0.44)	0.75 (0.43)	

Table 3: PSM-matched descriptive statistics by assigned EL status

language proficiency and student background covariate adjustments. These are the typical way that analysts attempt to account for bias in multiple regression methods. These result in a substantial attenuation of the relationship between EL-classification and teacher perceptions; though both relationships remain relatively large in magnitude and significant. Models 3 and 7 present our estimates of the relationship in our CEM-matched sample. The coefficients imply that an EL classification reduces teachers' perception of student ability in math by around one-twentieth of a standard deviation (though imprecisely estimated) and an imprecisely estimated null effect in reading. Finally, we present the results of our estimates in the PSM-matched sample in Models 4 and 8. Our estimate of the effect of classification on math and reading are slightly negative, but in both cases statistically equivalent to zero.

B6. We use a set of matching approaches to estimate the potentially causal effect of EL-classification on teachers' perceptions of kindergarten students' math and language abilities. We find that traditional statistical methods to adjust for background characteristics would misrepresent the effect of EL classification on teachers' perceptions of students' abilities. Whereas covariate-adjustment approaches would characterize EL classification as resulting in reduced teacher perceptions' of students' ability by 0.16 to 0.18 *SD*, matching approaches suggest that such classification does not change teacher perceptions for students in Kindergarten. Our preferred specification uses coarsened exact matching to identify a group of non-EL classified students who are equal in expectation to the EL-classified group. In these estimates, we cannot rule out that the effects of EL classification are indistinguishable from zero for teachers' perceptions of students' math and reading abilities. We argue our estimates should be understood as causal effects of EL classification if our assumption that students with the same language proficiency score are classified differently only based on idiosyncrasies of local policy.

Table 4: Comparison of estimates of EL classification on teacher perceptions

	Dependent variable:							
	Teacher Math Perception				Teacher Language Perception			
	OLS	OLS	CEM	PSM	OLS	OLS	CEM	PSM
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
EL-classified	-0.365^{***}	-0.160^{***}	0.004	-0.035	-0.453^{***}	-0.188^{***}	-0.064	-0.022
	(0.043)	(0.043)	(0.046)	(0.051)	(0.043)	(0.039)	(0.043)	(0.047)
Student Chars	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Observations	2,166	2,166	1,627	1,701	2,166	2,166	1,627	1,701
R^2	0.033	0.194	0.165	0.175	0.050	0.321	0.281	0.291

Notes: *p<0.05; **p<0.01; ***p<0.001. Table presents coefficients and standard errors (in parentheses).