capsee
CENTER FOR ANALYSIS OF POSTSECONDARY EDUCATION AND EMPLOYMENT

The Impact of Pell Grant Eligibility on Community College Students' Financial Aid Packages, Labor Supply, and Academic Outcomes

Appendices A, B, C, D, E, and F

A CAPSEE Working Paper

Rina Seung Eun Park
Community College Research Center
Teachers College, Columbia University

Judith Scott-Clayton
Community College Research Center
Teachers College, Columbia University
and NBER

March 2017

The authors' names are listed alphabetically. R. Park and J. Scott-Clayton contributed equally to this work. The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305C110011 to Teachers College, Columbia University. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education. The authors are grateful to the staff and administrators at the community college system that facilitated data access and interpretation.

For information about CAPSEE, visit capseecenter.org.

List of Appendices

Appendix A: Methods for Choosing Optimal Bandwidths **1**

A1. Leave-One-Out Cross Validation (Ludwig & Miller, 2005) 1
A2. Direct Plug-In Rule: Imbens and Kalyanaraman (2012) 1
A3. Second-Order Direct Plug-In Rule: Calonico, Cattaneo, and Titiunik (2014) 2
A4. Estimation of Bin Size 2
A5. Degree of Polynomial 2

Appendix B: Gerard, Rokkanen, and Rothe (2016) Bounds **4**

B1. Assumptions for GRR Bounds 4
B2. Implementing GRR Bounds 5

Appendix C: Sensitivity Checks **6**

Table C1: Testing for Continuity of Covariates 6
Table C2: Summary of Optimal Bandwidths 6
Table C3: Optimal Degree of Polynomial 7

Appendix D: Probability of Receiving Pell Grant (Left) and Average Amount of Pell Grant (Right) by EFC **8**

Figure D1 8

Appendix E: Baseline Covariance by EFC—Loan School Only **9**

Figure E1 9
Figure E2 10
Figure E3 11
Figure E4 12

Appendix F: Baseline Covariance by EFC—Continuous School Only **13**

Figure F1 13
Figure F2 14
Figure F3 15
Figure F4 16

Appendix A
Methods for Choosing Optimal Bandwidths

A1. Leave-One-Out Cross Validation (Ludwig & Miller, 2005)

Leave-one-out cross validation finds the optimal bandwidth by minimizing mean prediction errors—differences between the predicted value of *Y* and the actual value of *Y* for all observations *i* lying in the support, where the prediction comes from running each time a locally weighted regression using all the remaining observations on the same side of the cutoff (either left or right) but excluding observation *i.* Imbens and Lemieux (2008) and Imbens and Kalyanaraman (2012) suggest using local observations that are near to the cutoff. Following their suggestions, we trim observations on the extreme tails, taking only half the observations (those smaller than the median centered-EFC values) as the support on each side when running all regressions to predict Y. The objective is to minimize the average of prediction errors over all observations in the support. The cross-validation criterion is as follows:

$$h_{CV}^{opt} = argmin_h CV_Y^{.5}(h) = argmin_h [\frac{1}{N} \sum_{i:q_{X,.5,1} \leq X_i \leq q_{X,.5,r}} \left(Y_i - \hat{Y}(X_i)\right)^2]$$

A2. Direct Plug-In Rule: Imbens and Kalyanaraman (2012)

Imbens and Kalyanaraman (2012) (hereafter IK) propose to find an optimal bandwidth that minimizes the loss function, which is the mean of the squared prediction errors approaching from the left and right side of the cutoff (see below equation). To minimize this loss function, IK introduce a first-order approximation algorithm, referred to as the asymptotic mean squared error (AMSE). In mathematical notation, IK's optimal bandwidth algorithm calculates by using the following form:

$$h_{IK}^{opt} = argmin_h \mathrm{E}[((\hat{\mu}_+ - \mu_+)^2 - (\hat{\mu}_- - \mu_-))^2]$$

$$= C(k)[\frac{\sigma_+^2(x_0) + \sigma_-^2(x_0)}{f(x_0)\left(m_+^{(2)}(x_0)^2 - m_-^{(2)}(x_0)^2\right)}]^{\frac{1}{5}} N^{-\frac{1}{5}}$$

To obtain estimates of the unknown quantities in the formula, $\widehat{f(x_0)}$, $\hat{\sigma}$, and $\widehat{m}^{(2)}$, IK introduce a three-step algorithm that calculates consistent estimators of the parameters—first, calculate pilot bandwidth from Silverman's rule and estimate forcing variable density and conditional outcome variances at the cutoff; second, calculate preliminary bandwidths and use them to estimate second derivatives (curvature) at the cutoff; and third, add regularization terms to adjust for low precision for curvature estimators. IK's algorithm is one of plug-in rules to estimating optimal bandwidth.

A3. Second-Order Direct Plug-In Rule: Calonico, Cattaneo, and Titiunik (2014)

Calonico, Cattaneo, and Titiunik (2014) (hereafter CCT) minimize the same loss function as IK. The advancement in CCT is using bias-corrected estimators for unknown parameters to minimize the loss function by (1) constructing an alternative estimator for outcome variances that does not use a pilot bandwidth during the first step of IK's algorithm and (2) using consistent estimators for the preliminary bandwidths that are used to construct consistent population estimators during the second step of IK's algorithm (CCT).[1]

A4. Estimation of Bin Size

We follow McCrary's (2008) equation of bin size calculation $\hat{b} = 2\hat{\sigma}n^{-\frac{1}{2}}$ (where $\hat{\sigma}$ is the sample standard deviation of the running variable) and derive a \$100 centered-EFC bin size as optimal. We use \$100 centered-EFC values throughout the paper where bin size is required for estimation.

A5. Degree of Polynomial

Appendix Table C3 shows goodness-of-fit test results across different specifications—+/-\$2,000 bandwidth with covariates, without covariates, +/-\$1,000 bandwidth with covariates, and +/-\$4,000 bandwidth with covariates. Betas and standard errors represent coefficients for the treatment term and *p*-values are for the joint test on whether bin dummies are jointly equal to zero. We test goodness-of-fit up to polynomial of degree 4.[2] Polynomial of degree 0 function includes a constant and a treatment term where the coefficient of the treatment term is just a comparison between means of each side of the cutoff.

When we consider smaller or larger bandwidths (+/-\$1,000 or +/-\$4,000), the goodness-of-fit test suggests including higher order polynomials for our regression model. This is not surprising for a large bandwidth. However, it is surprising that for some variables the goodness-of-fit test suggests higher than order 4 polynomial even in a small bandwidth (+/-\$1,000). This may simply indicate that some of these variables are quite noisy right around the cutoff, rather than describing a true functional form. In our preferred model with a +/-\$2,000 bandwidth, zero polynomials are preferred for re-enrollment outcomes in Year 1 spring and fall, cumulative GPAs and earning outcomes, and earned any degree/certificate or transferred outcome in Year 3. For total loans received, total aid received, and full-time enrollment in Year 1 fall variables, the

[1] Please refer to CCT for further description of their algorithm.
[2] Gelman and Imbens (2014) suggest avoiding the use of very high degree of polynomial regressions in RD designs. Therefore, we focus up to polynomial degree 4. Nevertheless, we do not see any important changes in the goodness-of-fit significance level for degree 5 or 6 polynomials.

goodness-of-fit test recommends using a linear regression specification (degree 1). In general for +/-$2,000 bandwidth, we see drops in significance level at polynomial degree 1 and at polynomial degree 4. Since a lower degree of polynomial is preferred for local polynomial regressions (Gelman & Imbens, 2014), goodness-of-fit recommends degree 1 polynomial when using a +/-$2,000 bandwidth. We also see quite a drop in polynomial degree 1 and 2 for the +/-$1,000 bandwidth, while the +/-$4,000 bandwidth requires up to degree 3 to see a first drop in significance.

Appendix B
Gerard, Rokkanen, and Rothe (2016) Bounds

B1. Assumptions for GRR Bounds

Gerard, Rokkanen, and Rothe (2016) (hereafter GRR) require two additional assumptions regarding what they call the "selectors" (those students whose enrollment decisions shift as a result of their Pell eligibility): that the direction of selection is one-sided and that the conditional density is left-differentiable. Let $S_i \in \{0,1\}$ be indicators of sample selection (1 selector who changes enrollment decisions depending on funding package and 0 non-selectors).[3] We apply the usual RD assumptions to non-selectors—treatment probability is discontinuous at the cutoff, monotonicity, conditional expected potential outcome and conditional treatment probability is continuous at the cutoff, and the conditional running variable density for the non-selectors is differentiable and its derivative is continuous at the cutoff.

GRR assumption 1. Selectors are only located on one side of the cutoff (in our case, those who are eligible for Pell—i.e., the left side of the cutoff)

$$\Pr(\mathrm{X_i} < c | S_i = 1) = 1$$

GRR assumption 2. Selector's conditional density of a running variable is left-differentiable at the cutoff.

$$\mathrm{F}_{\mathrm{X|S=1}}(x) \text{ is left-differentiable in } x \text{ at } c$$

The added assumption implies that the left side of the cutoff is composed of only non-selectors (since all-selectors are opted out) while the right side of the cutoff has both selectors and non-selectors.[4] One of the major pitfalls of applying this commonly used "one-sided" assumption from prior literature (GRR, 2016; Lee, 2009; McCrary, 2008) to our sample selection problem is that it is not plausible to assume that the selectors all choose to opt out. There is no reason to believe that we will have one side with only the non-selectors.

---

[3] These selectors are equivalent to manipulators in GRR, as they are the ones who shift enrollment decisions (not to attend the community college system) because of the total funding package offered to those barely passing the Pell eligibility cutoff. We use the terms "selector vs. non-selector" rather than "manipulators vs. non-manipulators" because our density jump is a choice of enrollment selection and not a manipulation in running variable expected family contribution.

[4] The implication of this assumption in GRR differs slightly from that in our paper. GRR have higher density at the same side as both manipulators and non-manipulators (the treated side). Therefore, the one-sided assumption implies that the treated side consists of a mixture from both manipulators and non-manipulators while the non-treated side consists of only non-manipulators. On the other hand, in our case, selectors (manipulators in GRR) choose not to attend community college and therefore, decide to be out of our sample. Thus, we have lower density on the treated side (which includes only non-selectors) and higher density on the non-treated side (which includes a mixture of both selectors and non-selectors). This change of preference implies that the treated side will only consist of non-selectors while the control side has both selectors and non-selectors.

B2. Implementing GRR Bounds

The GRR bounding exercise takes the following steps. First, estimate the proportion of selectors (τ) by calculating the jump at the cutoff from the height of the density curve using local polynomial smoothing with rectangular kernel of degree 1 polynomial. Second, assuming selector as the best (worst) in outcomes, estimate the upper (lower) bound by the difference in expectation of outcome in the left and right side of the cutoff with post-trimmed smaller (larger) than τ $(1-\tau)$ quantiles (respectively).

Mathematically,

$$\Delta_0^U = \mathrm{E}[\mathrm{Y_i}|X_i = c^-] - E[Y_i|X_i = c^+, Y_i \geq Q_{Y|X}(\tau|c^+)]$$

$$\Delta_0^L = \mathrm{E}[\mathrm{Y_i}|X_i = c^-] - E[Y_i|X_i = c^+, Y_i \leq Q_{Y|X}(1-\tau|c^+)]$$

For discrete outcomes, trimming can be done by lower-coding (top-coding) the subtracted proportion, the proportion of those who are below (above) the conditional τ $(1-\tau)$ quantile – τ, for upper (lower) bounds (respectively).

In practice, GRR suggest using the polynomial truncation rule for trimming samples, which means using local polynomial approximations (we specify with rectangular kernel of degree 1 polynomial) to estimate conditional quantile functions to get τ $(1-\tau)$ quantile estimators. After trimming data using τ $(1-\tau)$ quantile estimators, we can estimate upper and lower bounds of differences in average outcome on each side of the cutoff by running a local polynomial regression (where we again use rectangular kernel of degree 1 polynomial). Because we have many outcomes of interest, we do not estimate standard errors for upper and lower bounds, which can be done by bootstrapping as introduced in GRR. In addition, further extension could be achieved by adding covariates from the start of calculating the density jump and throughout the steps to tighten the bounds. We do not employ this extension of adding covariates due to limitation in time and complexity rising from exploring numerous outcomes.

Appendix C
Sensitivity Checks

Table C1: Testing for Continuity of Covariates

	All Schools				Loan Schools			No-Loan Schools		
	Mean Outcomes	Basic 2000bw.			Mean Outcomes	Basic 2000bw.		Mean Outcomes	Basic 2000bw.	
Outcome	Just Above Cutoff	Coef.	(S.E.)		Just Above Cutoff	Coef.	(S.E.)	Just Above Cutoff	Coef.	(S.E.)
Female (%)	0.556	-0.037	(0.023)		0.550	-0.035	(0.027)	0.572	-0.045	(0.044)
Black (%)	0.242	-0.029	(0.020)		0.263	-0.015	(0.024)	0.178	-0.044	(0.033)
Hispanic (%)	0.071	-0.005	(0.011)		0.087	-0.001	(0.015)	0.022	-0.001	(0.013)
Asian (%)	0.050	0.000	(0.010)		0.064	0.003	(0.013)	0.004	0.004	(0.005)
American Indian (%)	0.004	0.001	(0.003)		0.004	0.001	(0.003)	0.004	0.001	(0.006)
Dual Enrollment (%)	0.227	0.044	(0.020)	**	0.182	0.028	(0.021)	0.368	0.047	(0.043)
Age	21.095	-0.056	(0.301)		21.196	-0.207	(0.338)	20.779	0.386	(0.639)
Income	$53,604	-$366	(739)		$53,688	-$600	(903)	$53,342	$88	(1241)
Depend	0.810	0.012	(0.018)		0.790	0.020	(0.021)	0.871	-0.020	(0.032)
Remedial Reading placement score	54.484	0.349	(1.807)		54.893	1.084	(2.113)	53.211	-1.410	(3.506)
Remedial Writing placement score	48.006	-0.167	(1.777)		48.999	0.525	(2.083)	44.910	-0.840	(3.406)
Remedial Math placement score	20.122	-0.564	(0.989)		20.177	-0.273	(1.194)	19.950	-1.128	(1.751)
Has Remedial Reading (%)	0.671	-0.007	(0.021)		0.676	0.000	(0.025)	0.656	-0.025	(0.042)
Has Remedial Writing (%)	0.681	-0.016	(0.021)		0.688	-0.010	(0.025)	0.660	-0.026	(0.042)
Has Remedial Math (%)	0.563	0.003	(0.023)		0.558	0.014	(0.026)	0.579	-0.025	(0.044)
Sample size	1,877	7,855			1,421	5,753		456	2,102	

Note. As a sensitivity check, we test continuity of covariates by replacing the dependent variable in the regression with covariates. We use the basic +/- $2,000 bandwidth with no covariate controls specification. Coefficients indicate beta values for indicator of treatment status (i.e., 1 if eligible for Pell and 0 otherwise).

*** $p < .01$, ** $p < .05$, * $p < .1$.

Table C2: Summary of Optimal Bandwidths

	(1) Cross-Validation	(2) Imbens & Kalyanaraman(2012)	(3) CCT(2014)
Mean	4,021	3,804	1,366
Minimum	802	1,638	718
25th quantile	3,167	3,173	1,188
50th quantile	4,613	3,563	1,307
75th quantile	5,270	4,218	1,467
Maximum	5,270	7,481	2,170

Note. Cross-validation and two plug-in estimators (IK and CCT) are estimated using the rdbwselect_2014 function in the Stata rdrobust package by Calonico, Cattaneo, Farrell, and Titiunik (2017).

Table C3: Optimal Degree of Polynomial

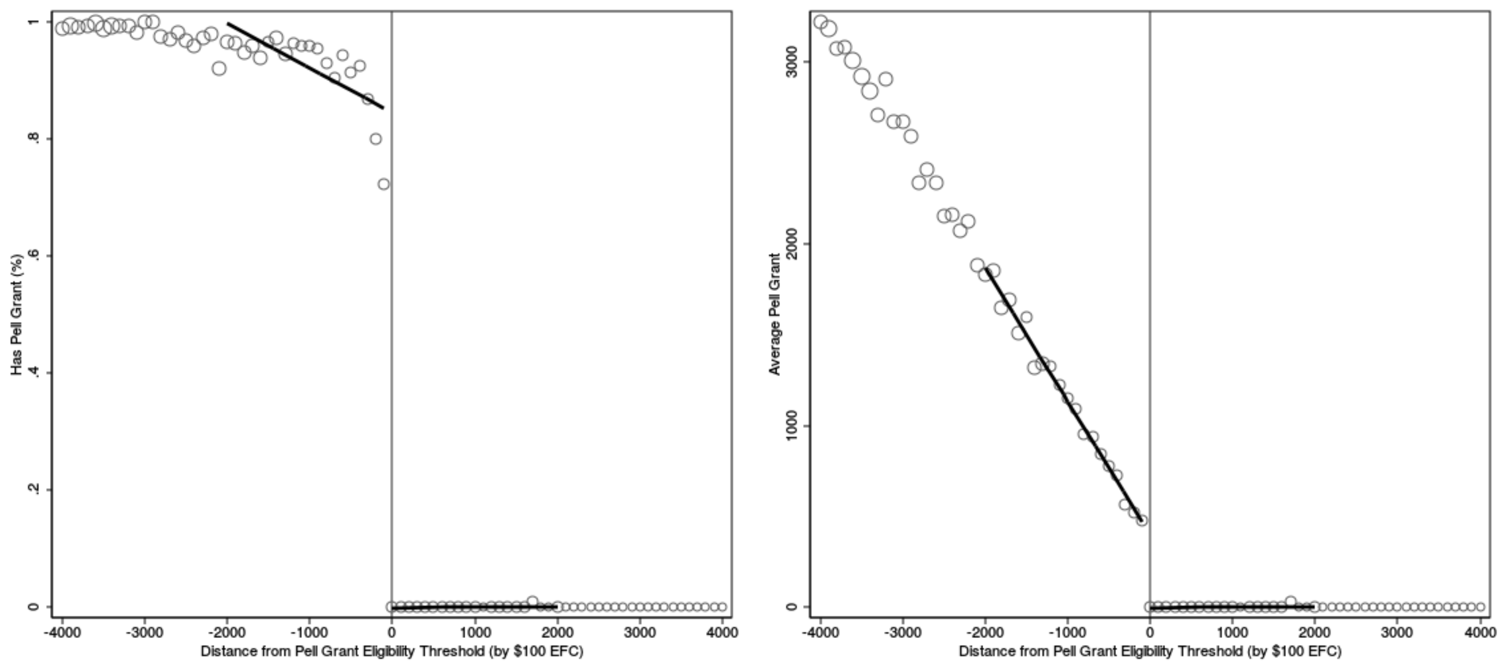
Outcome	(1) Poly. deg 0 (p-val)		(2) Poly. deg 1 (p-val)		(3) Poly. deg 2 (p-val)		(4) Poly. deg 3 (p-val)		(5) Poly. deg 4 (p-val)	
Amount of Pell received	0.000	***	0.027	**	0.030	**	0.033	**	0.038	**
Amount of Pell+State grants received	0.000	***	0.008	***	0.007	***	0.007	***	0.004	***
Amount of loans received	0.022	**	0.147		0.118		0.079	*	0.123	
Amount of total aid received	0.002	***	0.204		0.233		0.207		0.354	
Year 1 Outcomes										
Enrolled full-time, Year 1 Fall	0.047	**	0.336		0.375		0.414		0.273	
Re-enrolled, Year 1 Spring	0.805		0.718		0.642		0.649		0.696	
Enrolled full-time, Year 1 Spring	0.105		0.130		0.164		0.157		0.179	
Enrolled, Year 1 Summer	0.000	***	0.004	***	0.005	***	0.006	***	0.130	
Cum. GPA, End of Year	0.330		0.480		0.567		0.631		0.472	
Cum. Credits Completed, End of Year	0.009	***	0.011	**	0.012	**	0.014	**	0.006	***
Cum. Year 1 earnings (Q4-Q3)	0.320		0.270		0.249		0.241		0.276	
Year 2 Outcomes										
Re-enrolled, Year 2 Fall	0.632		0.555		0.609		0.698		0.653	
Enrolled full-time, Year 2 Fall	0.000	***	0.000	***	0.000	***	0.000	***	0.000	***
Re-enrolled, Year 2 Spring	0.000	***	0.000	***	0.000	***	0.000	***	0.000	***
Enrolled full-time, Year 2 Spring	0.000	***	0.000	***	0.000	***	0.000	***	0.032	**
Enrolled, Year 2 Summer	0.018	**	0.023	**	0.032	**	0.041	**	0.060	*
Cum. GPA, End of Year	0.317		0.621		0.585		0.602		0.513	
Cum. Credits Completed, End of Year	0.001	***	0.001	***	0.000	***	0.000	***	0.001	***
Cum. Year 2 earnings (Q4-Q3)	0.184		0.114		0.155		0.182		0.336	
End of Year 3 Attainment Outcomes										
Cum. GPA	0.295		0.674		0.650		0.621		0.550	
Cum. credits earned	0.000	***	0.000	***	0.000	***	0.000	***	0.000	***
Ever transferred to 4-Yr	0.073	*	0.095	*	0.146		0.302		0.468	
Earned any degree/cert	0.000	***	0.001	***	0.002	***	0.005	***	0.023	**
Earned any degree/cert or transferred	0.124		0.080	*	0.119		0.169		0.303	
Sample size	5,753		5,753		5,753		5,753		5,753	

Note. Samples are restricted to 2008–2010 fall entry cohort students who have filed FAFSA, for whom race/ethnicity is not missing, among those who are attending loan schools, and who are within a +/- $2,000 bandwidth. Beta and standard errors indicate coefficient for indicator variable of treatment status (i.e., eligible for Pell). Huber-White robust standard errors are in parentheses. *P*-values are for goodness-of-fit test, which tests for the null that all bin dummies (by Pell cutoff centered-EFC $100) are jointly equal to zero. Optimal degree polynomial is the degree where adding a higher order term no longer makes the bin dummies jointly significant. Polynomial degree 0 is comparing means left and right side of the cutoff with bin dummies. Column 1 specifies for polynomial degree 0, column 2 for degree 1, column 3 for degree 2, column 4 for degree 3, and column 5 for degree 4. All specifications control for cohort fixed effects.

*** $p < .01$, ** $p < .05$, * $p < .1$.

Appendix D
Probability of Receiving Pell Grant (Left) and Average Amount of Pell Grant (Right) by EFC

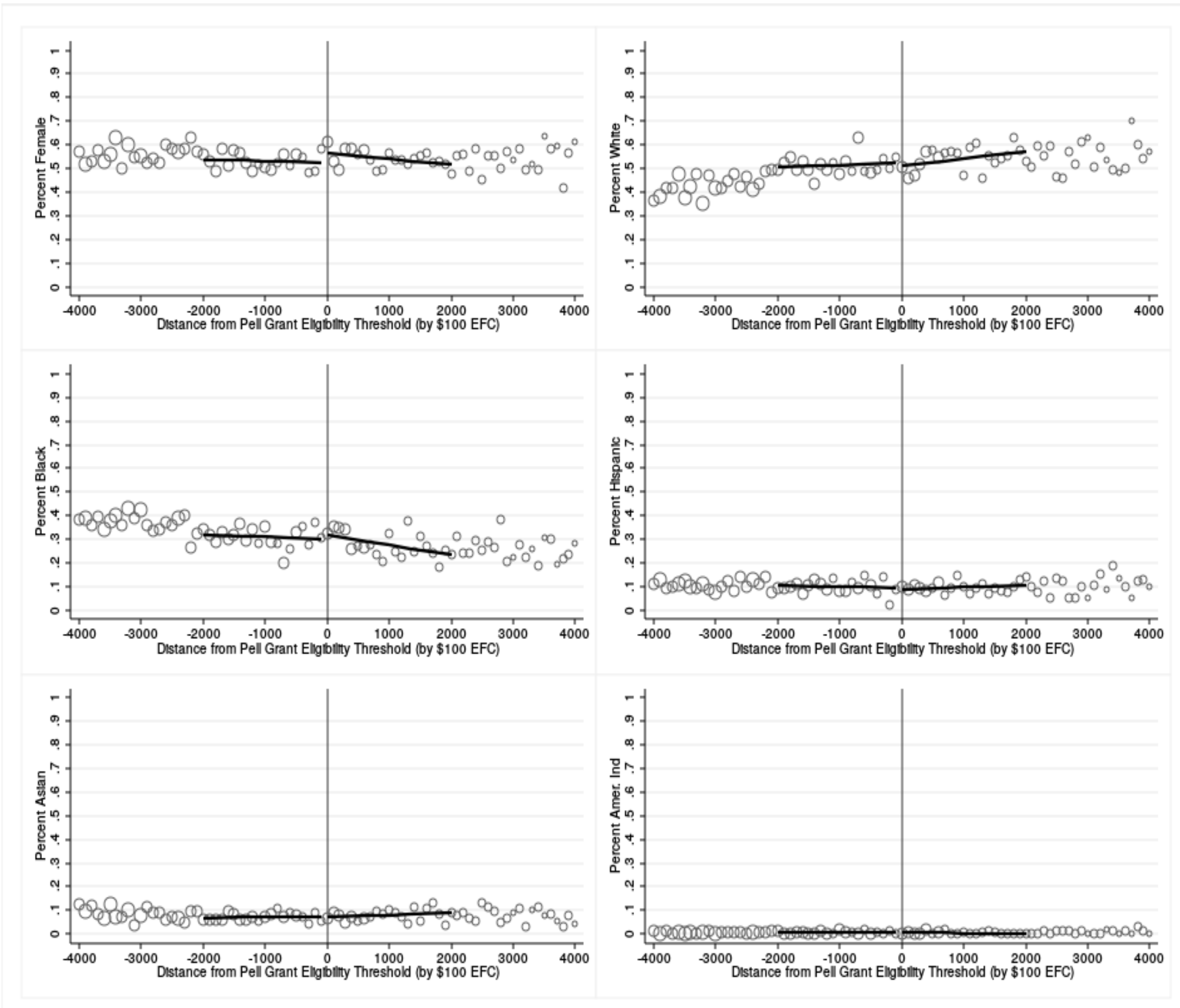
Figure D1



Note. Samples are restricted to 2008–2010 fall entry cohort students who have filed FAFSA, for whom race/ethnicity is not missing, and who are non-dual enrollees. Each point is a mean value of the outcome that falls within a bin size of $100 EFC. Graph shows only points that fall within +/- $4,000 bandwidth. Black line is a fitted line of mean points within a +/- $2,000 bandwidth.

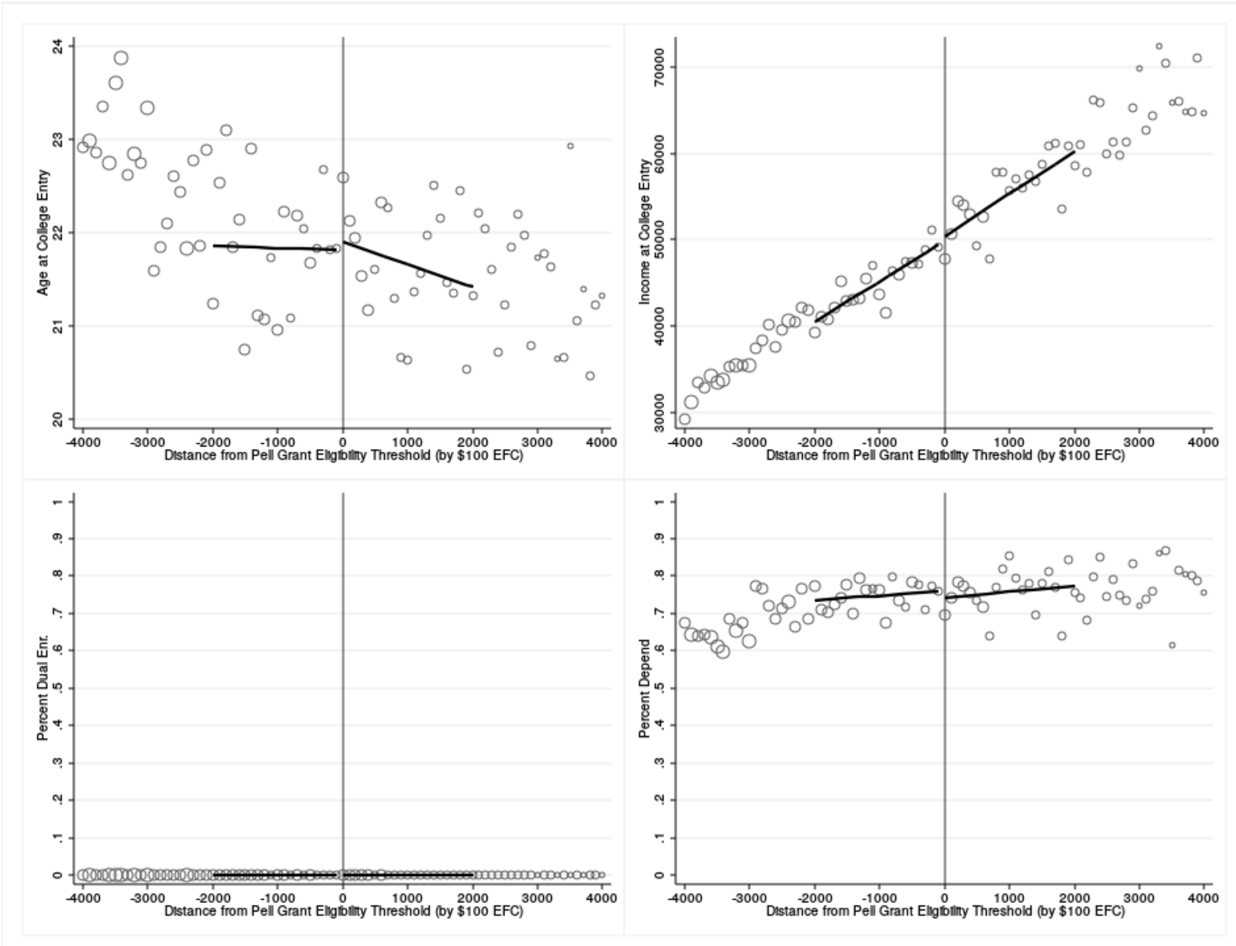
Appendix E
Baseline Covariance by EFC—Loan School Only

Figure E1



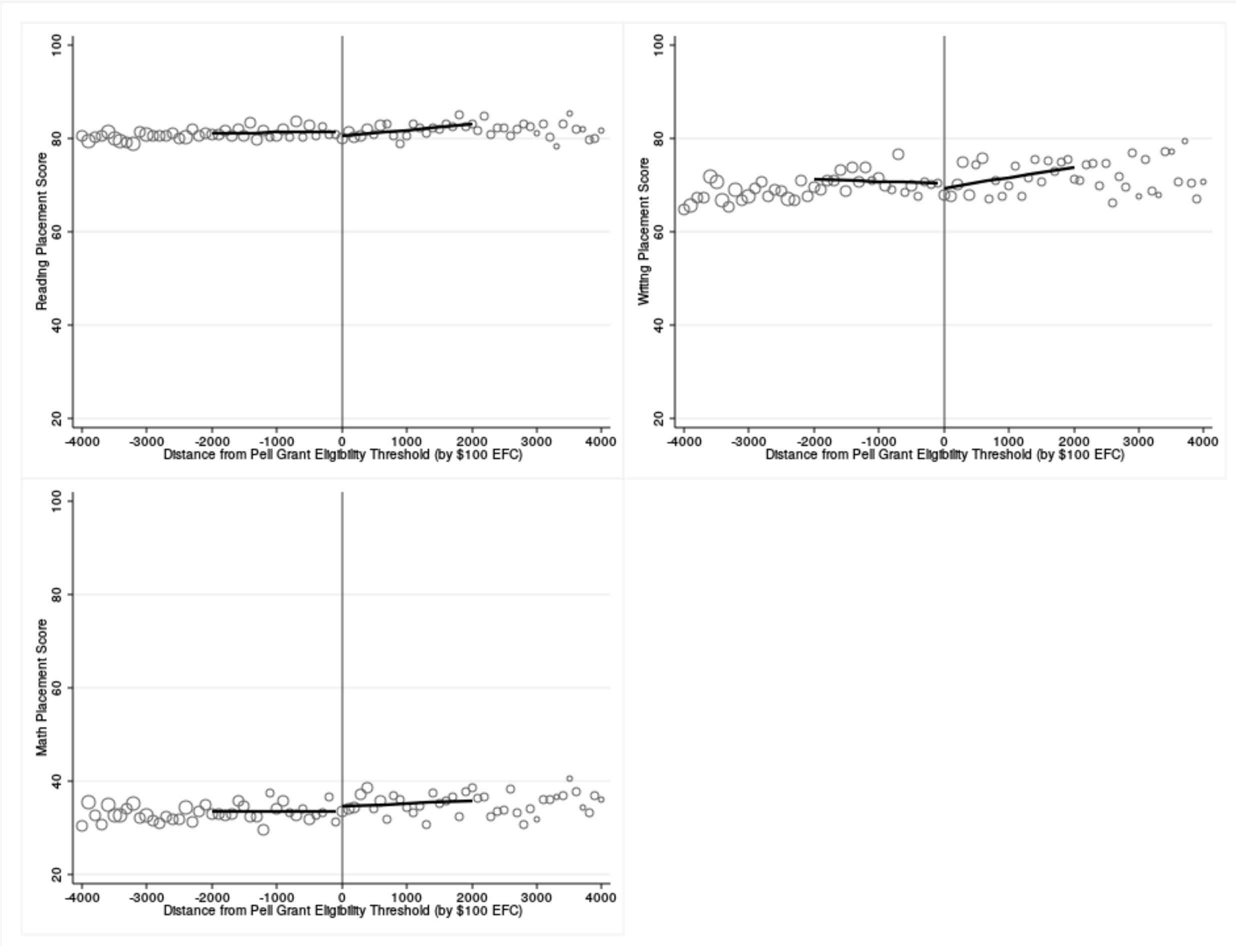
Note. Samples are restricted to 2008–2010 fall entry cohort students who have filed FAFSA, for whom race/ethinicty is not missing, who are non-dual enrollees, and who are attending loan schools only. Each point is a mean value of the outcome that falls within a bin size of $100 EFC. Graph shows only points that fall within a +/- $4,000 bandwidth. Black line is a fitted line of mean points within a +/- $2,000 bandwidth.

Figure E2



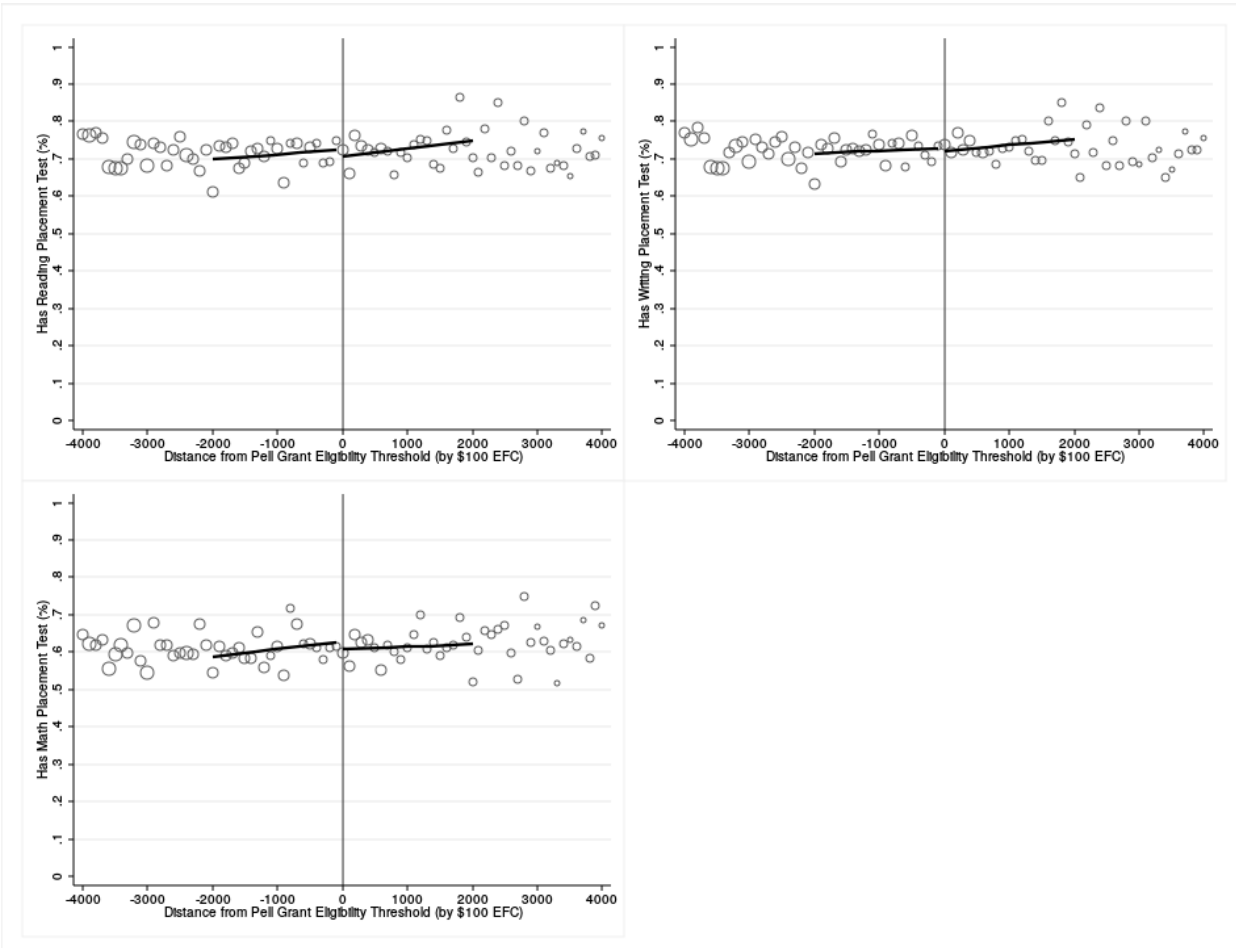
Note. Samples are restricted to 2008–2010 fall entry cohort students who have filed FAFSA, for whom race/ethnicity is not missing, who are non-dual enrollees, and who are attending loan schools only. Each point is a mean value of the outcome that falls within a bin size of $100 EFC. Graph shows only points that fall within a +/- $4,000 bandwidth. Black line is a fitted line of mean points within a +/- $2,000 bandwidth.

Figure E3



Note. Samples are restricted to 2008–2010 fall entry cohort students who have filed FAFSA, for whom race/ethnicity is not missing, who are non-dual enrollees, and who are attending loan schools only. Each point is a mean value of the outcome that falls within a bin size of $100 EFC. Graph shows only points that fall within a +/- $4,000 bandwidth. Black line is a fitted line of mean points within a +/- $2,000 bandwidth.

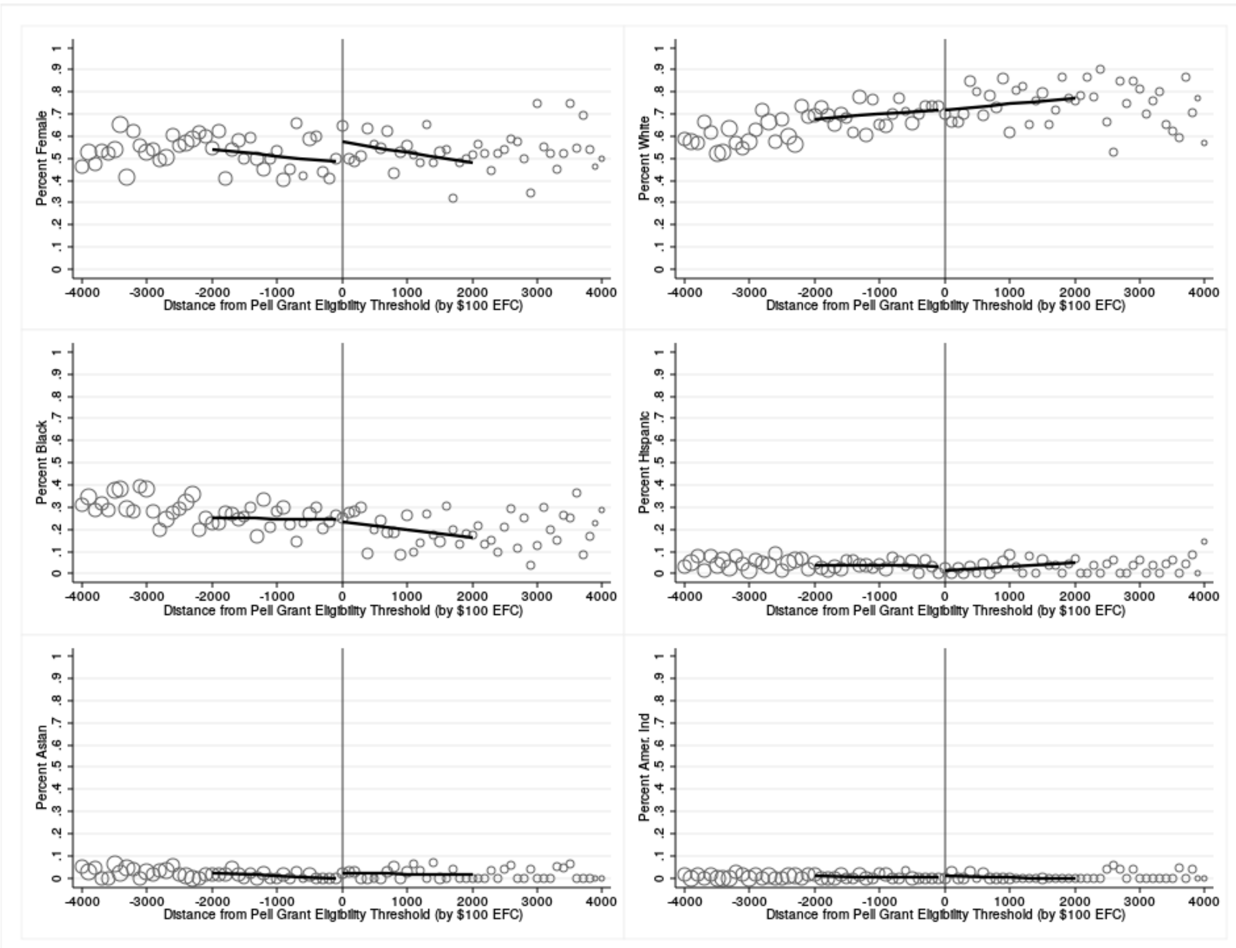
Figure E4



Note. Samples are restricted to 2008–2010 fall entry cohort students who have filed FAFSA, for whom race/ethnicity is not missing, who are non-dual enrollees, and who are attending loan schools only. Each point is a mean value of the outcome that falls within a bin size of $100 EFC. Graph shows only points that fall within a +/- $4,000 bandwidth. Black line is a fitted line of mean points within a +/- $2,000 bandwidth.

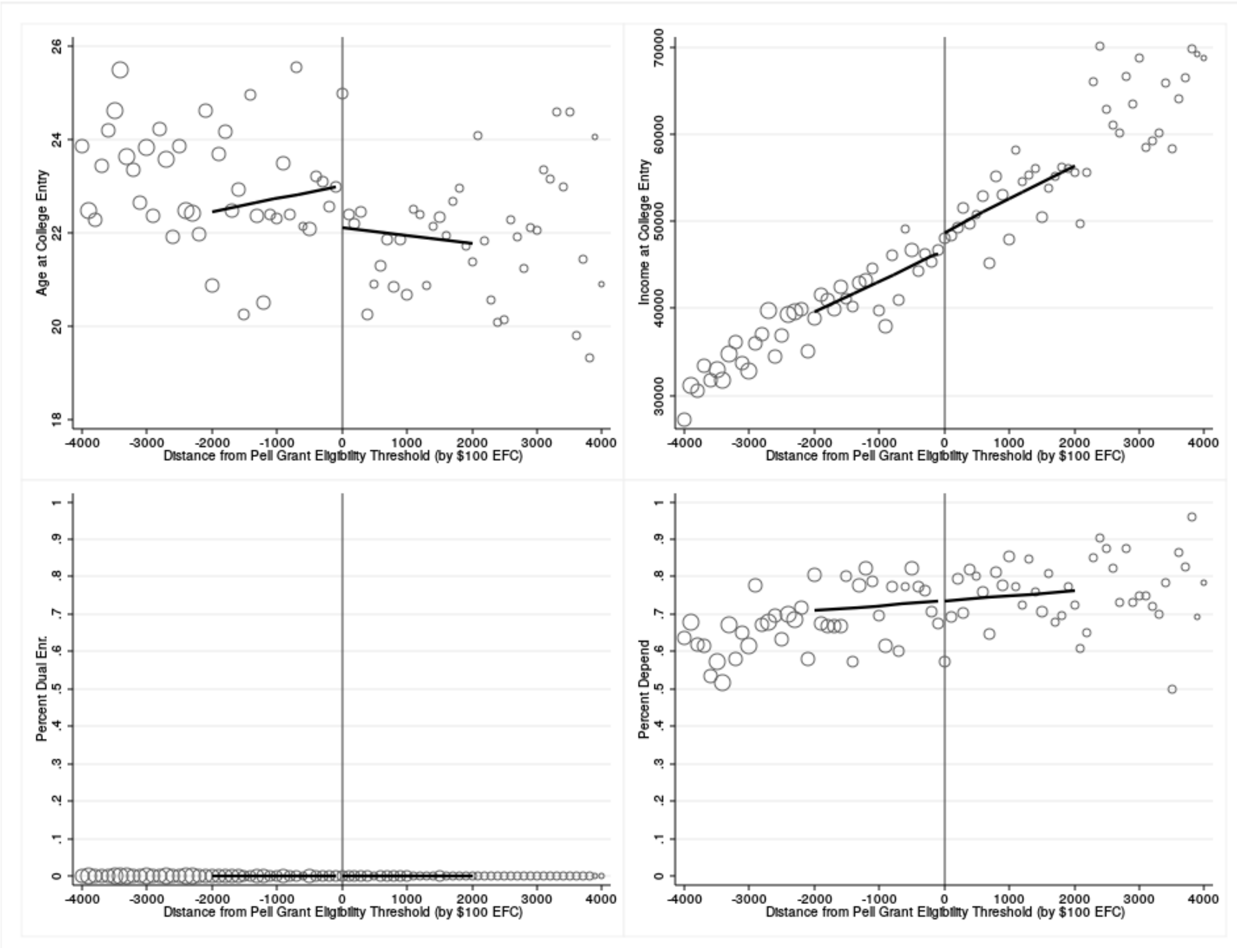
Appendix F
Baseline Covariance by EFC—Continuous School Only

Figure F1



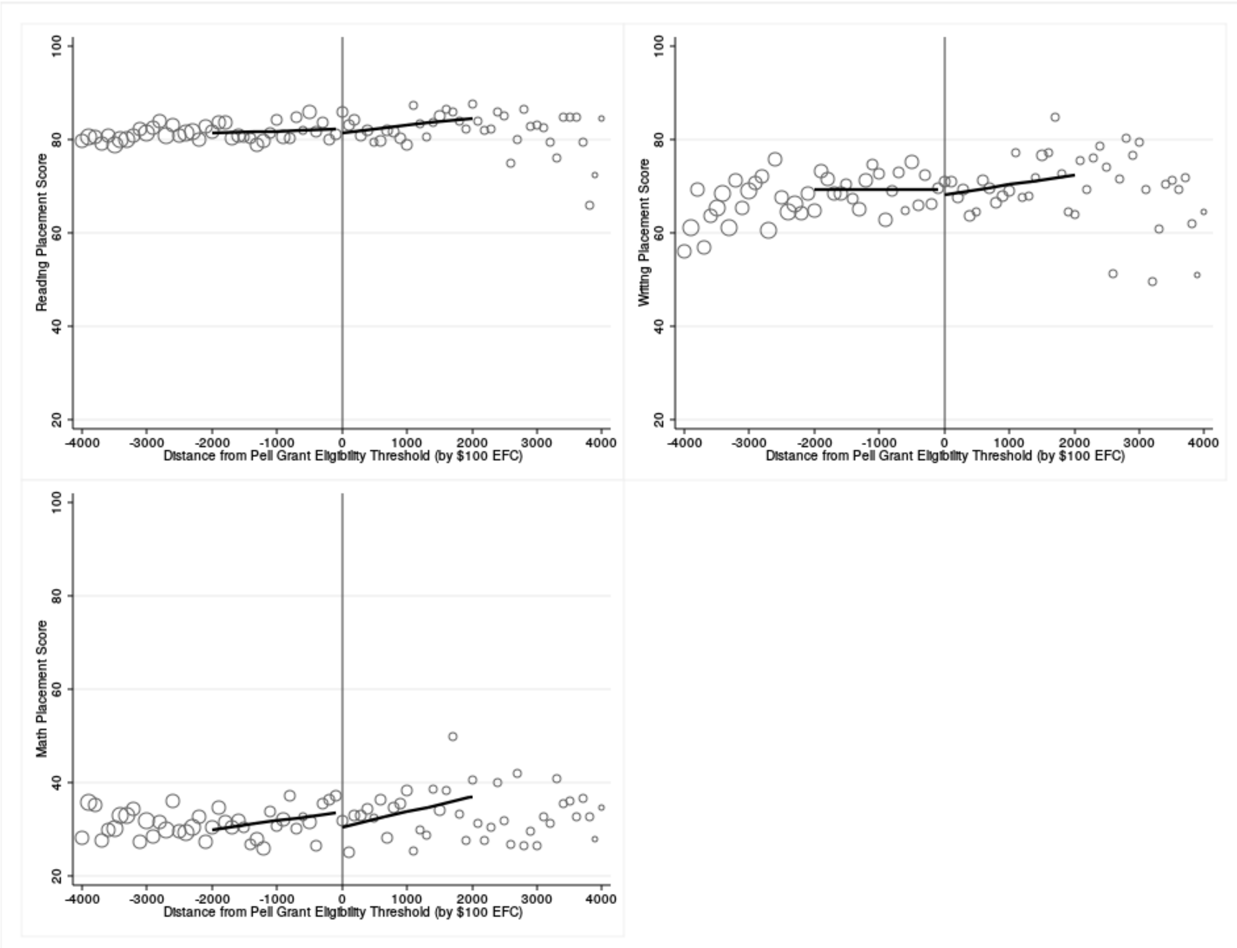
Note. Samples are restricted to 2008–2010 fall entry cohort students who have filed FAFSA, for whom race/ethnicity is not missing, who are non-dual enrollees, and who are attending loan schools that are defined as having a continuous density of observations around the cutoff as determined using the Calcagno and Long (2008) methodology. Each point is a mean value of the outcome that falls within a bin size of $100 EFC. Graph shows only points that fall within a +/- $4,000 bandwidth. Black line is a fitted line of mean points within a +/- $2,000 bandwidth.

Figure F2



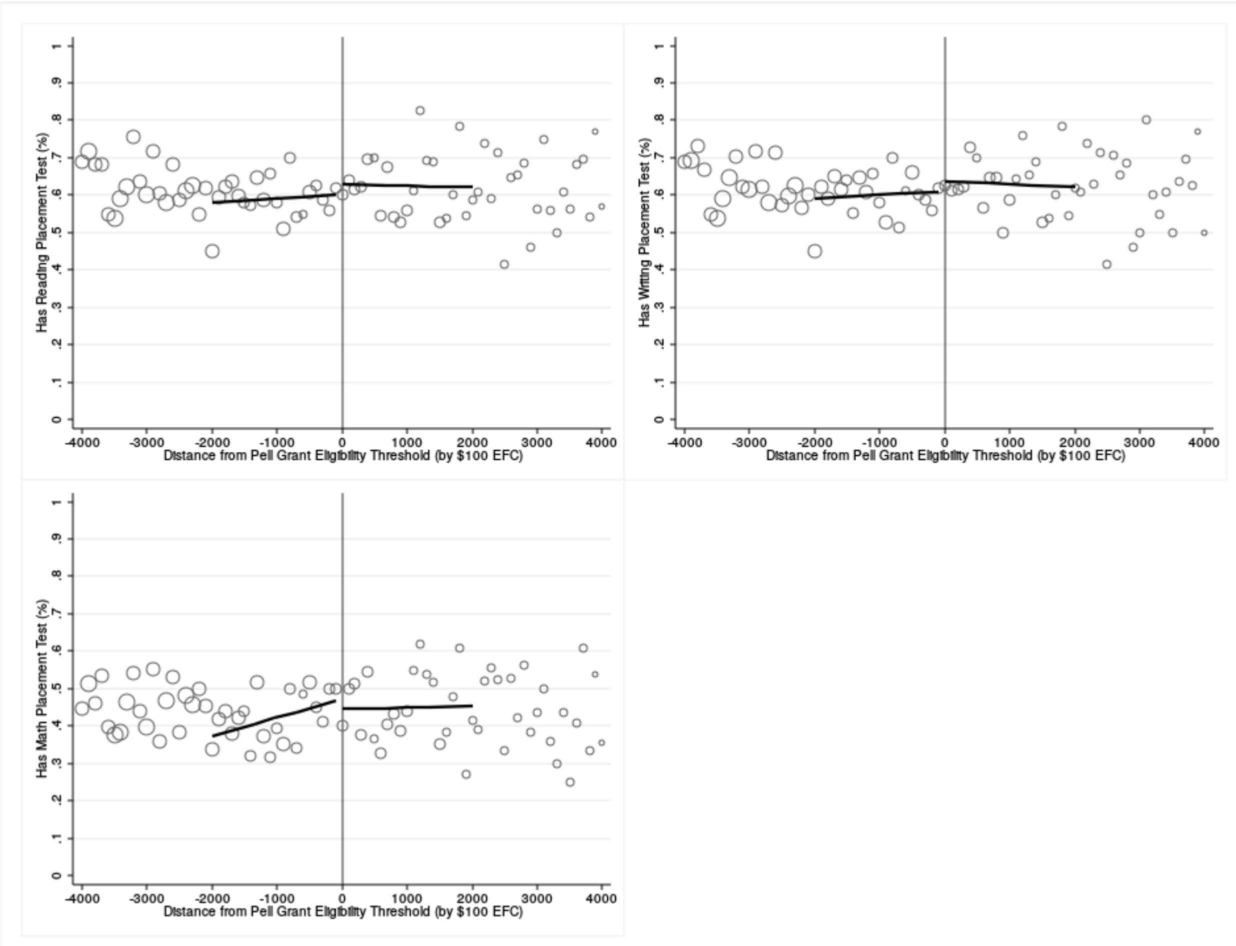
Note. Samples are restricted to 2008–2010 fall entry cohort students who have filed FAFSA, for whom race/ethnicity is not missing, who are non-dual enrollees, and who are attending loan schools defined as having a continuous density of observations around the cutoff, as determined using the Calcagno and Long (2008) methodology. Each point is a mean value of the outcome that falls within a bin size of $100 EFC. Graph shows only points that fall within a +/- $4,000 bandwidth. Black line is a fitted line of mean points within a +/- $2,000 bandwidth.

Figure F3



Note. Samples are restricted to 2008–2010 fall entry cohort students who have filed FAFSA, for whom race/ethnicity is not missing, who are non-dual enrollees, and who are attending loan schools defined as having a continuous density of observations around the cutoff, as determined using the Calcagno and Long (2008) methodology. Each point is a mean value of the outcome that falls within a bin size of $100 EFC. Graph shows only points that fall within a +/- $4,000 bandwidth. Black line is a fitted line of mean points within a +/- $2,000 bandwidth.

Figure F4



Note. Samples are restricted to 2008–2010 fall entry cohort students who have filed FAFSA, for whom race/ethnicity is not missing, who are non-dual enrollees, and who are attending loan schools defined as having a continuous density of observations around the cutoff, as determined using the Calcagno and Long (2008) methodology. Each point is a mean value of the outcome that falls within a bin size of $100 EFC. Graph shows only points that fall within a +/- $4,000 bandwidth. Black line is a fitted line of mean points within a +/- $2,000 bandwidth.