The Effects of Need-based Grant Aid on Long-Term College and Workforce Outcomes: Appendix

B. Heath Witzen

May 2019

1 Empirical Strategy

To estimate the effect of the EA Grant on various outcomes, a "fuzzy" regression discontinuity (RD) strategy is employed. Eligibility for the EA Grant is determined by a student's EFC, and this eligibility is used as an instrument for the amount of EA Grant received. Within a small enough bandwidth around the eligibility threshold and by controlling for a student's EFC, the being eligible is considered to be "as random" and eligibility is a valid instrument for EA Grant receipt.

To estimate the effect of the EA grant on various outcomes, the following method is used to estimate the following first and second stage equations for a student i of cohort t enrolled in institution j:

$$\begin{aligned} Y_{ijt} &= \beta EA\ Grant_{it} + \psi(\widetilde{X_{it}}) + \gamma_j + \gamma_t + \epsilon_{ijt} \\ EA\ Grant_{ijt} &= \delta \mathbb{1}\{\widetilde{X_{it}} < 0\} + \phi(\widetilde{X_{it}}) + \zeta_j + \zeta_t + \epsilon_{ijt} \end{aligned} \tag{1}$$

EA grant is used as a dependent variable in the first stage equation and EA grant is regarded continuous treatment. As observed by Lee and Lemieuxa [2010] the results of the fuzzy RD can be easily extended to this continuous treatment case. The variable $\widetilde{X_{it}}$ is the normalized distance from their EFC to the EFC of the cutoff, and the indicator function $\mathbb{1}\{\widetilde{X_{it}} < 0\}$ is equal to 1 if the EFC is below the threshold and equal to 0 if the EFC is above the threshold. The functions $\psi(\widetilde{X_{it}})$ and $\phi(\widetilde{X_{it}})$ are flexible functions of $\widetilde{X_{it}}$ as is typical in the RD literature. In the preferred specification, this is a degree one polynomial in $\widetilde{X_{it}}$, which is allowed to change in slope at the cutoff and vary by the entering cohort:

$$\phi(\widetilde{X_{it}}) = \zeta_1 \widetilde{X_{it}} + \zeta_2 \widetilde{X_{it}} \times \mathbb{1}\{\widetilde{X_{it}} < 0\} \tag{2}$$

In implementation, the functions are interacted with the cohort year allowing the slopes to differ by cohort year. In each estimation equation, fixed effects are included for the institution, cohort specific effects, and several control variables, such as race, gender, SAT math, and HSA scores. Equation 1 is estimated by two stage least squares (2SLS), using $\mathbb{1}\{\widetilde{X_{it}} < 0\}$ as an instrument for the amount of EA grant received.

The typical RD literature is followed by estimating these equations by choosing a subset of observations within a bandwidth of the EFC threshold. The Imbens and Kalyanaraman [2011] (hereafter IK) is used as the bandwidth measure, but the results are shown to be robust for a variety of bandwidths. This procedure yields a bandwidth of $3,500 EFC.

Since the threshold levels of EFC changed over time, this study also separately estimates the effect of $2,000 of EA Grant by the EFC threshold to test for whether there might be heterogeneous treatment effects by income, as income is highly related to EFC. The eight thresholds are split into "high" and "low", and the estimation of Equation 1 is used separately on each category. Because the IV procedure estimates a local average treatment effect, this process generates an estimate of the effect of EA Grant for those who received it around each level of EFC Threshold, and thus serves as a test for whether low and high income students react differently to EA Grant receipt. This estimation uses the same bandwidth as the pooled procedure.

The necessary assumption for the regression discontinuity approach to identify the causal effect of grant aid is that the expected potential outcomes in absence of the treatment are smooth through the threshold, or that there are no other variables that could influence the outcome (eg. persistence) that vary discretely at the threshold, only the EA Grant aid. Mathematically the necessary assumption is that conditional on the function $\phi(\widetilde{X_{it}})$, the mean of the unobserved variable ϵ_{ijt} does not vary discretely at the threshold or that

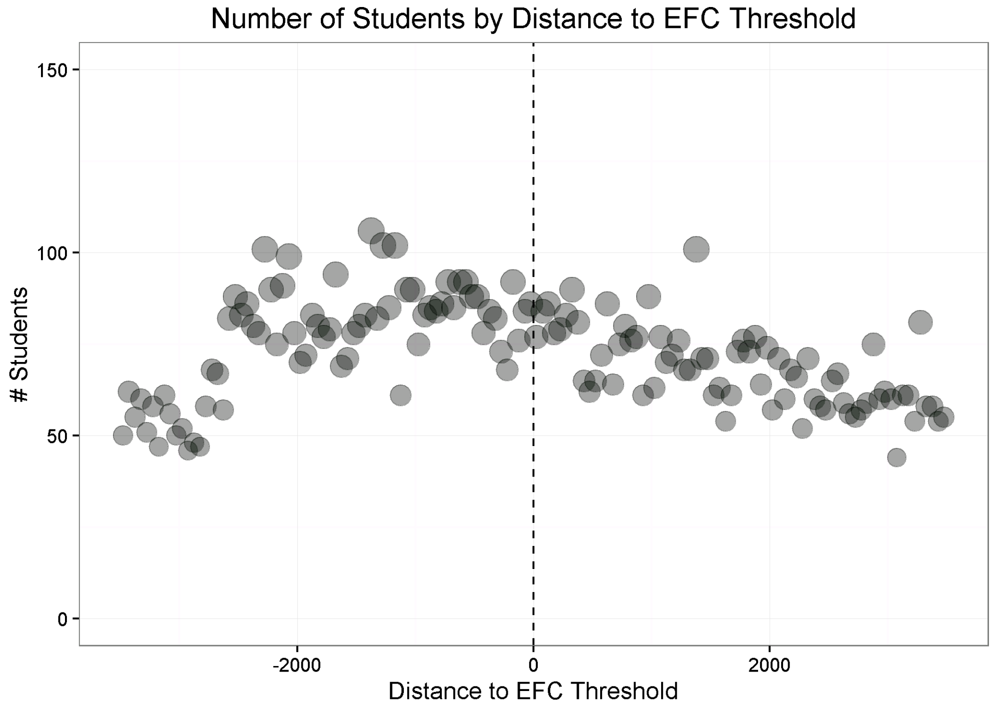
$$\lim_{\widetilde{X_{it}} \to 0^+} E[\epsilon_{ijt}|\widetilde{X_{it}}] = \lim_{\widetilde{X_{it}} \to 0^-} E[\epsilon_{ijt}|\widetilde{X_{it}}]$$

The necessary assumption is fundamentally untestable, but support for the validity can be shown by using standard diagnostics for RD designs. First, the density of students around the threshold can be shown to be smooth, suggesting that students cannot manipulate eligibility or that there are no initial enrollment effects of the EA Grant. Secondly, demographic characteristics can be observed to examine whether there is a discrete change in demographic characteristics at the threshold that would be possibly indicative of unobserved characteristics varying at the threshold.

2 Support for RD Assumptions

This section provides tests in support of the validity of the main RD Assumptions. Figure shows how the density of students evolves through the threshold and does not visually provide evidence of a change at the threshold. A McCrary (2008) test of the change in the density at the threshold produces a log-difference of -0.001 that is statistically indistinguishable from zero. This suggests that students are not sorting to one side of the EFC threshold or the other. This also serves as a test for an initial enrollment effect. This study can only observe a student's financial aid information if they enroll at a postsecondary institution. Therefore, if there was a significant initial enrollment effect, the density would be higher on the eligible side of the threshold compared to the ineligible side. This density plot and the McCrary test suggests that there is no initial enrollment effect. This is consistent with the fact that EA Grant students admitted off of the wait list may not find out until the Fall semester.

Figure 1: Density of Students at the Threshold



Note: Figure 1 shows the number of students within $50 EFC bins on each side of the EA Grant eligibility threshold.

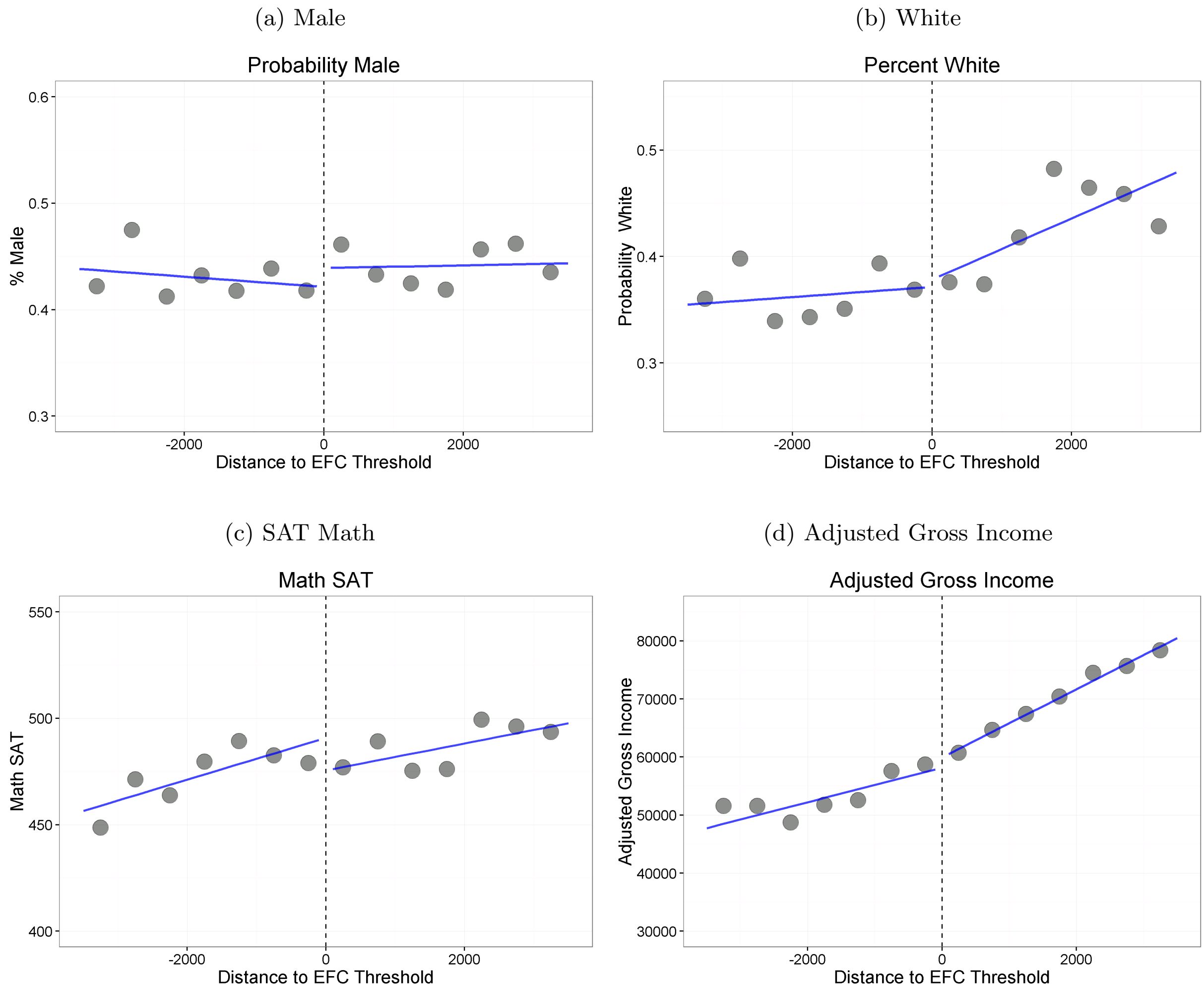
Table 1 shows how selected student characteristics vary at the eligibility threshold, and indicates no major changes in student demographic characteristics. Figure 2 plots the average student characteristics over the EFC distribution relative to the threshold. As demographic characteristics do not appear to change at the threshold, this provides additional support in favor of the required identification assumptions of the RD design.

Table 1: Checking for Demographic Changes at the EA Threshold

	Dependent variable:				
	Male	White	Hispanic	AGI	Math SAT
	(1)	(2)	(3)	(4)	(5)
EA Eligible	−0.01	0.002	0.004	958.00	8.11
	(0.02)	(0.02)	(0.01)	(722.79)	(6.53)
Dep. mean | Inelig.	0.44	0.43	0.06	69,700	486
Observations	10,227	10,227	10,227	10,227	10,227

Note: Table 1 displays the point estimates (and heteroskedasticity-robust standard errors in parentheses) of an RD regression using demographic characteristics as dependent variables. Each regression regresses the outcome on an indicator for having an eligible EFC, with a flexible function of EFC as a control. Estimates are obtained by a local linear regression with a rectangular kernel within the Imbens-Karyalanaraman (2011) bandwidth of $3,500 EFC. Refer to the methods section in the text for more details on the estimation. ***p<.01, **p<.05, $^{+}$p<.1

Figure 2: Changes in Demographic Thresholds

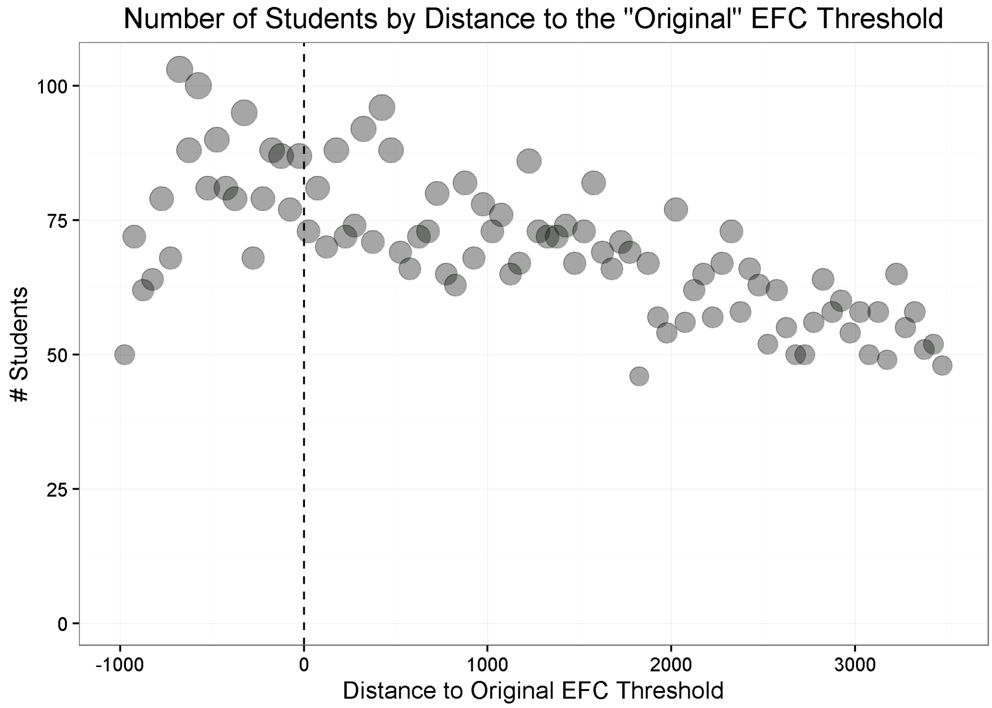


Note: Figure 2 shows the discontinuity in demographic characteristics at the eligibility threshold. The graph pools together all years of data. Gray dots represent the average EA Grant amount within \$500 EFC bins, while the solid line represents the estimated linear relationship estimated separately on each side of the threshold. Dollar amounts are in 2016 U.S. dollars.

3 Original Threshold

For the years 2010-2015, it is possible to use the same procedure using the initial cutoff, not the final cutoff after students have been admitted from the wait list. This provides both a robustness test and evidence that receiving notification of your award earlier is unlikely to drive enrollment. The density plot and associated McCrary density test suggests (though somewhat imprecisely) that students aren't sorting to one side of the initial eligibility threshold. Table 2 shows that there do not appear to be substantial demographic differences at the initial cutoff as well. Finally, Table 3 shows estimates for EA Grant receipt, institutional grant aid receipt, loan receipt, and persistence to the second year effects using the RD design with the initial threshold. There are no significant differences between those who are initially eligible and those who are not in EA Grant received, institutional grants received, or persistence to the second year. However, there is a significantly smaller amount of loans received by those who are initially eligible.

Figure 3: Density Plot at the Initial EFC Threshold



Note: Figure 3 shows the number of students, aggregated in $50 EFC bins on each side of the original EA Grant threshold applied before any students were admitted off of the wait list.

Table 2: Checking for Demographic Changes at the Original Threshold

	Dependent variable:				
	Male	White	Hispanic	AGI	Math SAT
	(1)	(2)	(3)	(4)	(5)
EA Eligible	0.01	0.01	−0.004	480.92	7.11
	(0.03)	(0.02)	(0.01)	(654.17)	(8.96)
Dep. mean | Inelig.	0.44	0.43	0.06	69,700	486
Observations	6,818	6,818	6,818	6,818	6,818

Note: Table 2 displays the point estimates (and heteroskedasticity-robust standard errors in parentheses) of an RD regression using demographic characteristics as dependent variables. Each regression regresses the outcome on an indicator for having an eligible EFC according to the original threshold, with a flexible function of EFC as a control. Estimates are obtained by a local linear regression with a rectangular kernel within the Imbens-Karyalanaraman (2011) bandwidth of $3,500 EFC. Refer to the methods section in the text for more details on the estimation. ***p<.01, **p<.05, $^{+}$p<.1

Table 3: Examining Original Cutoffs

	Dependent variable:			
	EA Grant	Inst. Grant	Loans	Pers.-Y2
	(1)	(2)	(3)	(4)
Immediately EA Eligible	−48	34	−525**	0.01
	(68)	(104)	(227)	(0.02)
Dep. mean | Inelig.	128	1,893	7,542	0.84
Observations	6,818	6,818	6,818	6,818

Note: Table 3 displays the point estimates (and heteroskedasticity-robust standard errors in parentheses) of an "fuzzy" RD regression using probabilities of financial aid and probability of persistence as dependent variables. The first stage regresses the amount of EA Grant receipt on an indicator for having an eligible EFC as of the original threshold. The second stage regresses the dependent variable on the estimated EA Grant aid. Estimates are obtained by a local linear regression with a rectangular kernel within the Imbens-Karyalanaraman (2011) bandwidth of $3,500 EFC. Refer to the methods section in the text for more details on the estimation. Column (1) displays the first stage estimates. Panel A uses EA Grant aid as the instrumented variable, while Panel B uses overall grant aid. ***p<.01, **p<.05, +p<.1

4 Robustness Tests

Tables 4, 6, and 5 show the robustness of the main estimated effects to the use of different bandwidths, functional forms, and polynomial of the EFC function.

Table 4 shows how financial aid, persistence, and workforce wage estimates differ by bandwidth. There are not meaningful differences in the estimated effects for financial aid or persistence outcomes by bandwidth. Workforce wages estimates do differ by bandwidth, but are quite imprecise with large standard errors.

Table 4: Robustness of Main Effects to Bandwidth Choice

	EFC Bandwidth:				
Dependent Var.	1000	2000	3000	4000	5000
EA Grant	1,514***	1,567***	1,629***	1,633***	1,679***
	(81)	(56)	(46)	(41)	(37)
Institution	-159	-163	-228**	-312***	-357***
	(194)	(135)	(108)	(96)	(86)
Total Loans	-780	-836**	-661**	-730***	-561**
	(514)	(354)	(286)	(259)	(235)
First Year Earn.	-497	-36	-63	-117	-34
	(319)	(222)	(177)	(157)	(142)
Pers.-2yr	0.081**	0.067***	0.045**	0.049***	0.047***
	(0.035)	(0.024)	(0.019)	(0.017)	(0.015)
Pers.-3yr	0.072+	0.064**	0.059***	0.057***	0.047***
	(0.04)	(0.027)	(0.022)	(0.02)	(0.018)
Pers.-4yr	0.048	0.041	0.043+	0.036	0.032
	(0.045)	(0.031)	(0.026)	(0.023)	(0.02)
Degree-5yrs	0.065	0.039	0.043	0.025	0.023
	(0.059)	(0.041)	(0.033)	(0.029)	(0.026)
Earn-5yrs	1,053	1,011	541	-360	-494
	(2,820)	(1,966)	(1,605)	(1,376)	(1,200)
Earn-6yrs	7,557+	4,478	3,600	2,692	1,128
	(4,366)	(3,052)	(2,376)	(1,990)	(1,663)
Earn-7yrs	4,907	4,787	6,521+	5,443+	1,316
	(7,140)	(4,723)	(3,529)	(2,925)	(2,465)
N	3,203	6,269	8,980	11,106	13,123

Note: Table 4 displays the point estimates (and heteroskedasticity-robust standard errors in parentheses) of an "fuzzy" RD on each dependent variable, varied by chosen bandwidth. The first stage regresses the amount of EA Grant receipt on an indicator for having an eligible EFC. The second stage regresses the dependent variable on the estimated EA Grant aid. Estimates are obtained by a local linear regression with a rectangular kernel within the Imbens-Karyalanaraman (2011) bandwidth of \$3,500 EFC. Refer to the methods section in the text for more details on the estimation. Each cell is a separate estimation, with the dependent variable as the row and the columns representing the chosen bandwidth. ***p<.01, **p<.05, $^{+}$p<.1

Table 5 shows the estimated effect by the polynomial of the EFC function used in the 2SLS estimation. EA

Grant and persistence effects do not appear to differ much by polynomial order. Institutional grants do seem to be sensitive to polynomial choice, as do workforce wages.

Table 5: Robustness of Main Effects to the Degree of Polynomial

	Polynomial Order:		
Dependent Var.	1	2	3
EA Grant	1,620.94***	1,555.77***	1,448.42***
	(43.00)	(64.40)	(85.83)
Institution	-253.19***	-12.13	-153.11
	(81.16)	(121.71)	(162.28)
Total Loans	-563.03***	-515.37	-620.27
	(218.07)	(326.92)	(435.91)
Pers.-2yr	0.036**	0.051**	0.077***
	(0.014)	(0.021)	(0.028)
Pers.-3yr	0.051***	0.045+	0.064+
	(0.017)	(0.025)	(0.033)
Pers.-4yr	0.032+	0.037	0.031
	(0.019)	(0.029)	(0.038)
Degree	0.022	0.022	0.046
	(0.023)	(0.035)	(0.046)
Earn.-5yrs	269.06	1,007.18	833.97
	(1,076.98)	(1,602.63)	(2,141.44)
Earn.-6yrs	2,828.99+	2,711.99	4,607.00
	(1,509.57)	(2,252.60)	(3,023.51)
Earn.-7yrs	5,246.07**	2,583.64	1,581.87
	(2,248.96)	(3,369.49)	(4,483.18)

Note: Table 5 displays the point estimates (and heteroskedasticity-robust standard errors in parentheses) of an "fuzzy" RD on each dependent variable, varied by polynomial of the flexible function. The first stage regresses the amount of EA Grant receipt on an indicator for having an eligible EFC. The second stage regresses the dependent variable on the estimated EA Grant aid. Estimates are obtained by a local linear regression with a rectangular kernel within the Imbens-Karyalanaraman (2011) bandwidth of $3,500 EFC. Refer to the methods section in the text for more details on the estimation. Each cell is a separate estimation, with the dependent variable as the row and the columns representing the chosen order of polynomial. ***p<.01, **p<.05, +p<.1

Lastly, Table 6 shows how robust the main estimates are to the functional form and controls used in the estimating equation. This table largely shows that the estimates are not sensitive to the fixed effects and controls used in the estimation.

Table 6: Robustness of Main Effects to Functional Form and Controls

	Functional Form:			
EA Grant	1,638.36***	1,624.97***	1,620.99***	1,620.94***
	(42.13)	(43.32)	(43.03)	(43.00)
Institution	-151.85	-263.54**	-307.89***	-312.40***
	(120.93)	(123.73)	(102.04)	(101.82)
Total Loans	-1,088.88***	-638.67**	-718.72***	-694.70**
	(267.72)	(277.23)	(272.52)	(270.60)
First Year Earn.	-44.17	-127.34	-115.02	-118.24
	(160.10)	(164.86)	(167.27)	(166.43)
Pers.-2yr	0.038**	0.045**	0.045**	0.045**
	(0.017)	(0.018)	(0.018)	(0.018)
Pers.-3yr	0.063***	0.066***	0.063***	0.063***
	(0.02)	(0.021)	(0.021)	(0.021)
Pers.-4yr	0.045+	0.044+	0.04+	0.04+
	(0.023)	(0.024)	(0.024)	(0.024)
Degree-5yrs	0.043	0.043	0.03	0.029
	(0.031)	(0.032)	(0.031)	(0.031)
Earn-5yrs	-184.37	-213.42	282.32	366.92
	(1,413.10)	(1,494.57)	(1,481.71)	(1,480.32)
Earn-6yrs	4,040.71+	4,051.21+	4,108.01+	4,011.37+
	(2,256.90)	(2,257.16)	(2,197.06)	(2,193.04)
Earn-7yrs	8,058.45**	8,022.73**	7,849.77**	7,480.64**
	(3,365.56)	(3,359.04)	(3,313.63)	(3,300.94)
N	10,227	10,227	10,227	10,227
Year interaction?	No	Yes	Yes	Yes
Institution FE?	No	No	Yes	Yes
Demographic Vars.?	No	No	No	Yes

Note: Table 6 displays the point estimates (and heteroskedasticity-robust standard errors in parentheses) of an "fuzzy" RD on each dependent variable, varied by chosen functional form and controls. The first stage regresses the amount of EA Grant receipt on an indicator for having an eligible EFC. The second stage regresses the dependent variable on the estimated EA Grant aid. Estimates are obtained by a local linear regression with a rectangular kernel within the Imbens-Karyalanaraman (2011) bandwidth of $3,500 EFC. Refer to the methods section in the text for more details on the estimation. Each cell is a separate estimation, with the dependent variable as the row and the columns representing the chosen functional form. ***p<.01, **p<.05, +p<.1

Table 7 provides estimates of the main effects excluding two years where the EA Grant threshold was very close to the Pell Grant threshold to demonstrate how the main effects are not likely to be due to differential Pell receipt.

Table 7: Without Years 2011 and 2013 and Away from the Pell Threshold

	Dependent variable:										
	EA Grant	Pell	Inst.	Loans	Pers.-Y2	Pers.-Y3	Pers.-Y4	Degree	Earn-5ys	Earn-6yrs	Earn-7yrs
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Original	1,625***										
EA Cutoff	(60)										
EA Grant (per $2,000)		−0.27	−438***	−618**	0.06***	0.07***	0.05*	0.04	647	5,681**	7,481**
		(26)	(119)	(314)	(0.02)	(0.02)	(0.03)	(0.04)	(1,720)	(2,844)	(3,282)
Dep. mean | Inelig.	128	844	1,893	7,542	0.84	0.75	0.71	0.62	18,178	21,211	24,794
Observations	7,370	7,147	7,147	7,147	7,147	7,147	5,545	3,128	3,128	1,951	1,951

Note: Table 7 displays the point estimates (and heteroskedasticity-robust standard errors in parentheses) of an “fuzzy” RD regression using fiancial aid, probabilities of persistence and degree receipt, and earnings as dependent variables. The first stage regresses the amount of EA Grant receipt on an indicator for having an eligible EFC. The second stage regresses the dependent variable on the estimated EA Grant aid. Estimates are obtained by a local linear regression with a rectangular kernel within the Imbens-Karyalanaraman (2011) bandwidth of $3,500 EFC. Refer to the methods section in the text for more details on the estimation. Column (1) displays the first stage estimates. Years 2011 and 2013, in which the EFC threshold is close to the Pell threshold have been removed. ***p<.01, **p<.05, $^{+}$p<.1

4.1 Dynamic Regression Discontinuity -Additional Robustness Test

One aspect of the design of the EA grant program that affects the estimated parameter of the RD strategy is that students who are ineligible for an EA Grant may subsequently become eligible for EA Grants in later years. An example of this would be a student who is just ineligible for the EA Grant in their first year, given their EFC. In their second year, the EA Grant threshold is more lenient (higher) than in their first year, and their EFC now qualifies for an EA Grant. Once a student becomes eligible, they may renew the EA Grant in any subsequent years they are enrolled. Subsequent eligibility for students who were ineligible in their first year may be gained by a change in family circumstances or by the threshold rising in a subsequent year.[1] Students who enter in 2013 and are just ineligible by EFC are likely to be eligible in their next year (2014) for a renewable EA grant.

A consequence of a policy where ineligible students can subsequently become eligible is that it changes the interpretation of the "control" group (those just ineligible in their first year) compared to the treatment group. Some ineligible students can receive treatment in their later years. For the sake of notation, I will consider year 0 to be a student's first year, and year 1 the second year, etc. If I want to consider the effect of EA Grant eligibility in the first year on enrollment in year 1, then an RD estimation of:

$$y_{1i} = \beta g_{1i} + f(\tilde{EFC}_{1i}) + \epsilon_i \tag{3}$$

identifies the effect β. However, this β has a less intuitive interpretation. It is now the effect of becoming eligible in the first year minus a treatment effect for the proportion of the control group that become eligible in the second year multiplied by the probability that a student in the control group received treatment. This effect is smaller than the effect of becoming eligible for the EA Grant in the first year versus a counterfactual in which the student never receives a decrease in price due to the EA Grant.

To address the unique nature of the EA Grant program and estimate a policy-relevant treatment effect, the effect of becoming eligible for the EA Grant in a student's first year, I adapt a "dynamic" RDD model of **?**. In this model, eligibility for an increase in grant aid can affect the probability of grant aid increases in the future years. I consider a treatment indicator $g_{i,t}$ that is equal to 1 if a student i is permanently eligible for an extra amount of grant aid up to $3,000 in year t and zero otherwise. To represent the renewable nature of the EA Grant, $g_{i,t}$ is an indicator for receiving an increase in grant aid in year t that decreases tuition in each subsequent year after t. As an outcome variable, I consider y_{it} as an indicator for whether a student is enrolled in year t. If the direct effect of receiving an increase in grant aid in year $t-\tau$ on enrollment in year y_{it} depends on only the number of years since the increase in grant aid, then y_{it} can be written as the sum of grant aid changes in each previous year:

$$y_{it} = \sum_{\tau=0}^{\infty} g_{i,t-\tau}\beta_\tau^D + \epsilon_{it} \tag{4}$$

or as the sum of the partial effects of the complete history of increases in grant aid. The coefficient β_τ^D is the direct effect of a increase τ years prior to t on y_{it} absent any other increases in grant aid. This is denoted by D, and is the effect of receiving the increase in aid, holding constant all other grant aid amounts.

The direct effects are policy relevant when considering the effect of grant aid. For example, a policymaker might want to know what is the effect of providing a student an extra $3,000 in grant aid, beginning in year 1, where the aid is renewable and the student knows that it is guaranteed in all years, on the probability of enrolling in year 3. In a model using y_{i3} as the dependent variable, this would be the effect β_3^D.

An RD regression like that of Equation **??** could be performed on an indicator for receiving a grant increase τ years earlier. Such a regression would take the form:

$$y_{it} = g_{i,t-\tau}\beta_\tau^T + f(EFC) + e_{it} \tag{5}$$

[1] In Figure **??**, which shows the EA Grant thresholds in each year, it is possible to see an example of the latter situation.

However, as explained above, the identified effect includes the direct effect of a grant increase plus the effects on the probability of future grant aid receipt. This regression identifies an "total", T, effect, which includes the effect on future grant aid increases.[2] If it is possible that receiving a increase in grant aid changes the probability of receiving a permanent increase in grant aid in the future, then β_1^T will not equal β_1^D. The total effect then is a combination of the direct effect of receiving an increase in a given year and the probability that subsequent treatments will be received. Equation 6, shows how, if the probability of receiving a permanent increase in the future depends on receiving an increase in a prior year, then the total effect of β_τ equals:

$$\beta_\tau^T \equiv \frac{dy_{it}}{dg_{i,t-\tau}} = \frac{\partial y_{it}}{\partial g_{i,t-\tau}} + \sum_{h=1}^{\tau} \Big(\frac{\partial y_{it}}{\partial g_{i,t-\tau+h}} * \frac{dg_{i,t-\tau+h}}{dg_{i,t-\tau}} \Big) \tag{6}$$

$$= \beta_\tau^D + \sum_{h=1}^{\tau} \beta_{\tau-h}^D \pi_h \tag{7}$$

where π_h equals the change in probability of a grant increase in period $t-\tau+h$ due to receiving a grant increase in $t-\tau+h$.

As a concrete example, the effect of receipt in a student's first year on persistence to the second year can be written as:

$$\beta_1^T = \beta_1^D + \beta_0^D \pi_1 \tag{8}$$

$$= \frac{\partial y_{i1}}{\partial g_{i,0}} + \frac{\partial y_{i1}}{\partial g_{i,1}} * \frac{dg_{i,1}}{dg_{i,0}} \tag{9}$$

or the effect of receiving the grant in the first year on enrollment in the second year plus the effect of receiving the grant in the second year on enrollment in the second year multiplied by the change in the probability of receiving the grant in the second year after receiving the grant in the first year.

In the case of the EA Grant program, students who are ineligible in year one may receive the EA Grant if they qualify in a later year. This means that $\pi_h < 0$, and assuming that the direct effects are positive, then the $\beta_\tau^D > \beta_\tau^T$. Another way to think about this is to consider the treatment and control group when the treatment is receiving an EA Grant in the first year. If the probability of receiving an increase in year 2 is affected, then the total effect incorporates the fact that some of the initially eligible students received treatment. For many grant program design questions, this is unlikely to be policy relevant.

Following the method of **?**, I implement a recursive estimator to use total effects to estimate the direct effects of receiving increases in grant aid. Using the RD estimator, as shown in Equation 3, the total effect of receiving an increase in grant aid in the first year is identified for any year. Using a recursive method, each total effect can be written as:

$$\beta_0^T = \beta_0^D \tag{10}$$

$$\beta_1^T = \beta_1^D + \pi_1 \beta_0^D \tag{11}$$

$$\beta_2^T = \beta_2^D + \pi_2 \beta_1^D + \pi_1 \beta_0^D \tag{12}$$

$$\beta_3^T = \beta_3^D + \pi_3 \beta_2^D + \pi_2 \beta_1^D + \pi_1 \beta_0^D \tag{13}$$

To estimate β_0^T, I use an RDD regression of enrollment in year 1 on EA Grant eligibility in year 1. For β_1^T, I use an RDD regression of enrollment in year 2 on EA grant eligibility in year 1, and so on. The π effects are similarly intent to treat effects, or the overall effect of receiving a permanent increase in aid in a given year due to a change in receiving grant aid in the first year. For example π_1 can be identified by a regression of

[2] In **?**, the direct effects are called "treatment on the treated" and total effects are called "intent to treat" effects, mirroring the language used in an instrumental variables setting. Here I use "direct" and "total" to prevent confusion, because I will estimate the RD using a "fuzzy" design and do not want to confuse two types of fuzziness.

the indicator for receiving a grant aid increase on EA Grant eligibility in year 1, and all other πs estimated in a similar manner. Once the total effects and πs are estimated, then the estimates of the direct effects can be derived, and standard errors for the D estimates can be obtained by the delta method.

Table 8: Main Academic Effects- Receipt of EA Grant

	Persistence to the/Graduation in:			
	2nd	3rd	4th	Grad-5yrs
A. Without dynamic estimation				
Effect of eligiblity	0.037**	0.051**	0.033+	0.023
	(0.015)	(0.017)	(0.02)	(0.025)
Effect of EA Grant receipt	0.065**	0.091**	0.059+	0.043
	(0.026)	(0.03)	(0.035)	(0.047)
95% ci	(0.015 , 0.115)	(0.032 , 0.15)	(-0.01 , 0.128)	(-0.048 , 0.135)
B. With dynamic estimation				
Effect of eligiblity	0.037**	0.058**	0.043+	0.043
	(0.015)	(0.019)	(0.023)	(0.023)
Effect of EA Grant receipt	0.065**	0.102**	0.076+	0.065
	(0.026)	(0.034)	(0.042)	(0.053)
95% ci	(0.015 , 0.115)	(0.036 , 0.168)	(-0.005 , 0.158)	(-0.027 , 0.169)
Dep. mean | Inelig.	0.84	0.75	0.71	0.62
Observations	10,227	10,227	8,625	6,208

Note: Table 8 displays the point estimates and standard errors (in parentheses) of an RD on each dependent variable. Estimates are obtained by a local linear regression with a rectangular kernel within the Imbens-Karyalanaraman (2011) bandwidth of $3,500 EFC. Panel A computes the estimates without the dynamic framework of section **??**, while Panel B incorporates dynamics in the estimation. The first stage for the effects on the 4th year and 5th year graduation are .557 and .523, respectively. Refer to the methods section in the text for more details on the estimation. Each cell is a separate estimation, with the dependent variable as the row and the columns representing the chosen order of polynomial. ***p<.01, **p<.05, +p<.1

Table 9: Academic Effects: Low Versus High Income

	Persistence to the/Graduation in:			
	2nd	3rd	4th	Grad-5yrs
A. Low Income				
first stage	0.614***	0.614***	0.614***	0.562***
	(0.016)	(0.016)	(0.019)	(0.024)
Effect of eligiblity	0.06**	0.07**	0.043	0.043
	(0.018)	(0.024)	(0.035)	(0.035)
Effect of EA receipt	0.097**	0.113**	0.07	0.124
	(0.029)	(0.04)	(0.056)	(0.078)
95 ci	(0.04 , 0.155)	(0.035 , 0.192)	(-0.04 , 0.181)	(-0.083 , 0.277)
B. High Income				
first stage	0.552***	0.552***	0.552***	0.532***
	(0.019)	(0.019)	(0.019)	(0.022)
Effect of eligiblity	0.017	0.023	0.032	0.032
	(0.018)	(0.023)	(0.024)	(0.024)
Effect of EA receipt	0.031	0.041	0.058	0.001
	(0.033)	(0.041)	(0.044)	(0.054)
95 ci	(-0.034 , 0.096)	(-0.039 , 0.121)	(-0.029 , 0.145)	(-0.047 , 0.106)

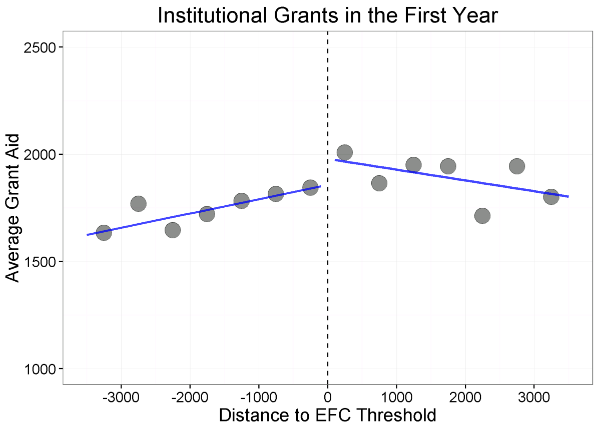
Note: Table 9 displays the point estimates and standard errors (in parentheses) of an RD on each dependent variable, and indicator of enrollment in a given year and enrolled in a STEM major. Estimates are obtained by a local linear regression with a rectangular kernel within the Imbens-Karyalanaraman (2011) bandwidth of $3,500 EFC. Each panel uses the dynamic framework of Section **??**. Panel A uses students that face a "low" EFC threshold for eligibility in their first year and Panel B shows estimates for students who face a "high" threshold. Refer to the methods section in the text for more details on the estimation. Each cell is a separate estimation, with the dependent variable as the row and the columns representing the chosen order of polynomial. ***p<.01, **p<.05, $^{+}$p<.1

5 Additional Tables and Figures

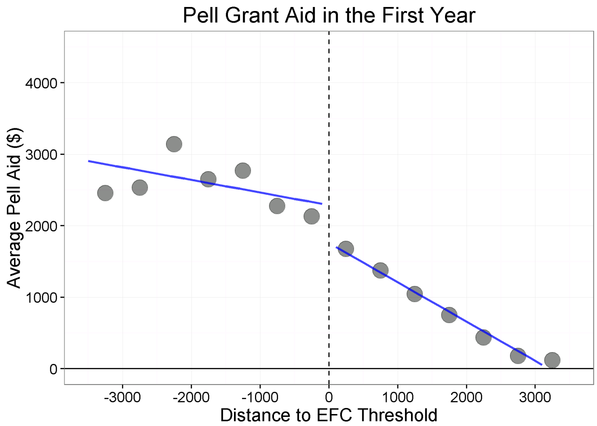
Figures 4, 5 and 6 provide visual representations of the RD effects found in Tables 2, 4 and 5 of the report.

Figure 4: Changes in Financial Aid Variables at the Threshold

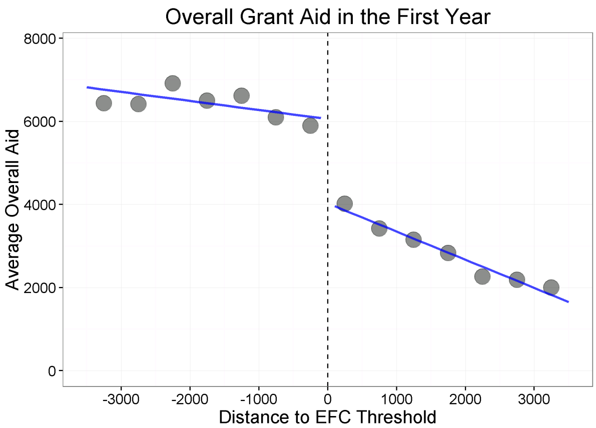
(a) Institutional Grants



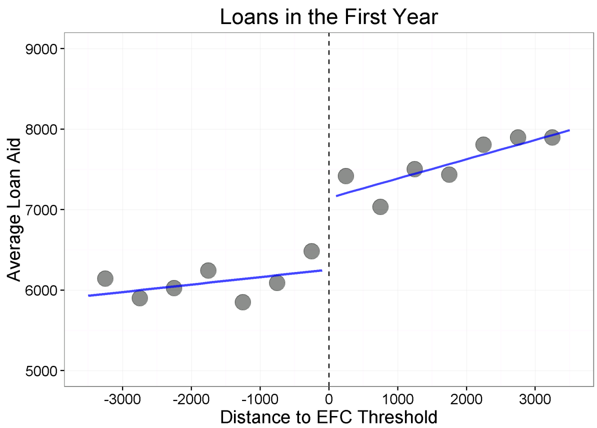
(b) Pell Grants



(c) Overall Grant Aid



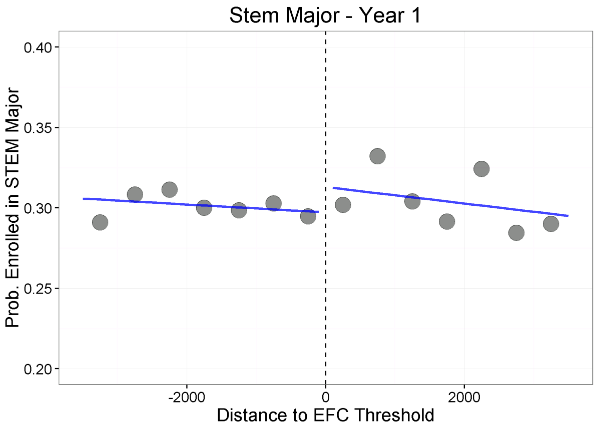
(d) Loans



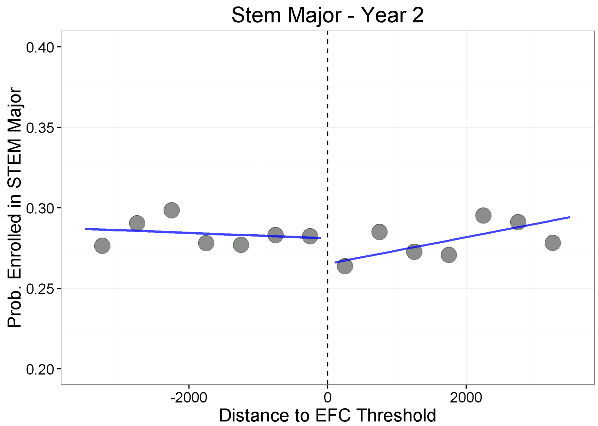
Note: Figure 4 shows the discontinuity in the types of financial aid at the eligibility threshold. The graph pools together all years of data. Gray dots represent the average EA Grant amount within $500 EFC bins, while the solid line represents the estimated linear relationship estimated separately on each side of the threshold. Dollar amounts are in 2016 U.S. dollars.

Figure 5: Probability of Being Enrolled in a STEM Major

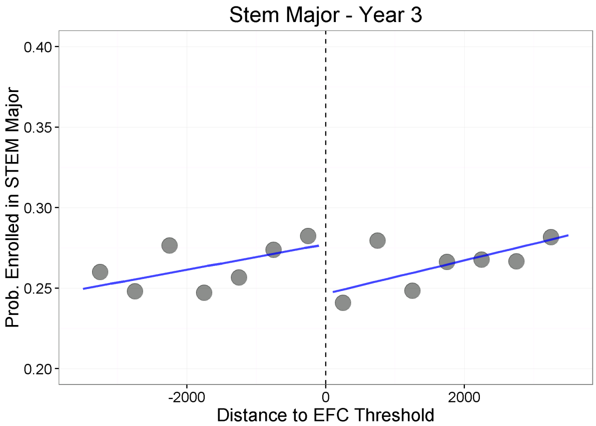
(a) Year 1



(b) Year 2



(c) Year 3



(d) Year 4

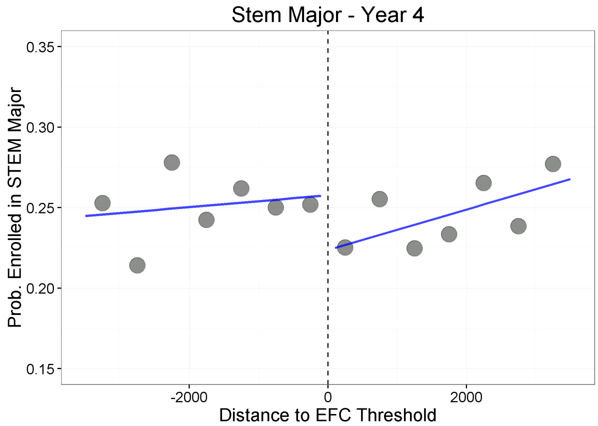
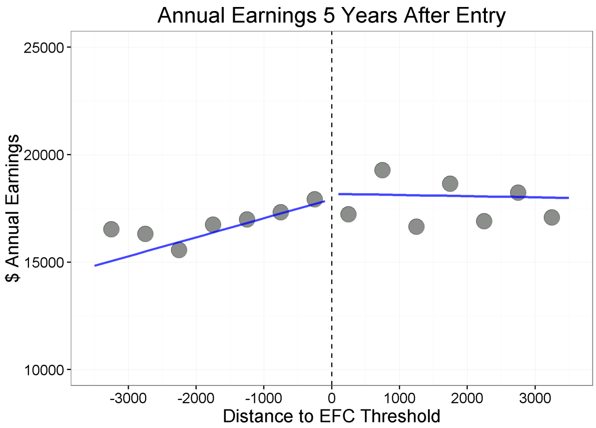


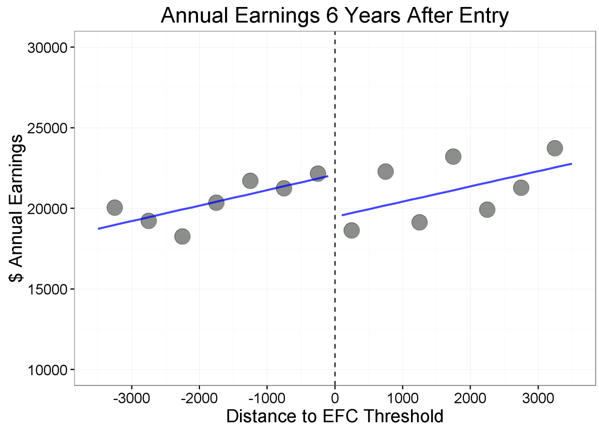
Figure 5 shows the discontinuity in the probability of being enrolled in a STEM major at the eligibility threshold. Persistence and graduation are measured by the probability of the outcome variable. The graph pools together all years of data. Gray dots represent the average in $500 EFC bins, while the solid line represents the estimated linear relationship estimated separately on each side of the threshold.

Figure 6: Changes in Workforce Wages at the Threshold

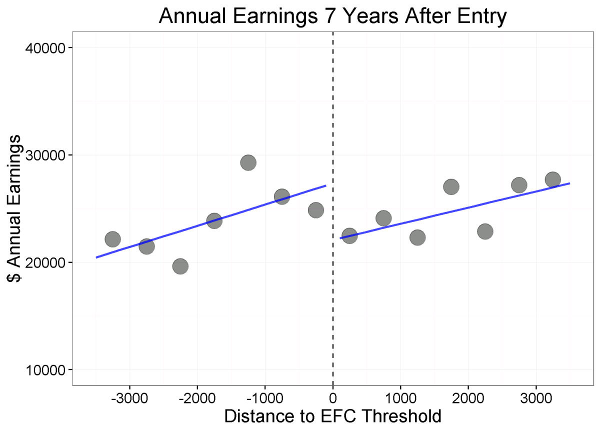
(a) 5th Year



(b) 6th Year



(c) 7th Year



Note: Figure 6 shows the discontinuity in earnings after a given number of years at the eligibility threshold. The graph pools together all years of data. Gray dots represent the average EA Grant amount within $500 EFC bins, while the solid line represents the estimated linear relationship estimated separately on each side of the threshold. Dollar amounts are in 2016 U.S. dollars.

6 Other Tests of Program Design

This section provides additional tests of program design. Table 10 estimates the effect of beginning to receive the EA Grant in a student's second year on persistence to the third year. This process mirrors the main estimation but uses the studen'ts second year EFC and the threshold that a student would have faced in their second year instrument for receipt beginning in year two on persistence to year 3. Table 11, compares the estimates for the EA Grant program to that of the Pell Grant program. Pell Grant receipt is likewise determined on a scale according to EFC, and there is an upper limit (between $5,000 EFC and $6,000 EFC over this time period) above which students can no longer receive Pell Grant. The same RD estimation can thus be used on the Pell Grant and the estimate can be compared to the EA Grant.

In Table 10, it is possible to see in column (1) that student's who are just eligible are more likely to begin receiving the EA Grant beginning in year 2, thus producing a relevant first stage for the estimation. Columns (2) and (3) show the estimated reduced form and IV effects of EA Grant eligibility or receipt, respectively, beginning in year 2. This estimation suggests that those who begin receiving the EA Grant in year two do not experience positive effects, however, the standard errors are fairly large and the estimates imprecise.

Table 10: Receiving Grant Aid beginning Year 2

	Dependent variable:		
	Rec. Year 2	Pers-Y3	
	(1)	(2)	(3)
EA Eligible	0.37***	−0.005	
	(0.02)	(0.02)	
EA Grant receipt			−0.01
			(0.04)
Observations	9,363	9,363	9,363

Note: Table 10 displays the point estimates (and heteroskedasticity-robust standard errors in parentheses) of an "fuzzy" RD regression using persistence to the second year as a dependent variable. The first stage regresses the an indicator for beginning to receive EA Grant in year 2 on an indicator for having an eligible EFC in year 2. The second stage regresses the dependent variable on the estimated EA Grant aid. Estimates are obtained by a local linear regression with a rectangular kernel within the Imbens-Karyalanaraman (2011) bandwidth of $3,500 EFC. Refer to the methods section in the text for more details on the estimation. ***p<.01, **p<.05, $^{+}$p<.1

Table 11 shows the first stage of the Pell and EA grant estimations in columns (1) and (2) respectively, showing how eligible students in each case are more likely to receive Pell Grant and EA Grant aid. Columns (3) and (4) show the estimated IV effects of $1,000 of Pell Grant and $1,000 of EA Grant Aid on persistence to year 2, respectively. The EA Grant effect is positive and significant, while the Pell Grant effect is slightly negative and statistically insignificant. These results suggest that the EA Grant has larger effects than the Pell Grant, but should be treated as suggestive, as the confidence intervals overlap, and the effects cannot be distinguished.

Table 11: EA Vs. Pell

	Dependent variable:			
	Pell Y1	EA Year 1	Pers-Y2	
	(1)	(2)	(3)	(4)
Pell Eligible	535.51***			
	(14.69)			
EA Eligible		1,677.38***		
		(37.94)		
Pell Receipt			−0.005	
			(0.03)	
EA Grant Receipt				0.02***
				(0.01)
Observations	12,283	13,123	12,283	13,123

Note: Table 11 displays the point estimates (and heteroskedasticity-robust standard errors in parentheses) of an "fuzzy" RD regression using persistence to the second year as a dependent variable. The first stage regresses the average amount Pell Grant or EA Grant on an indicator for eligibility for Pell and EA Grant respectively. The second stage regresses the dependent variable on the estimated grant aid. Estimates are obtained by a local linear regression with a rectangular kernel within the Imbens-Karyalanaraman (2011) bandwidth of \$3,500 EFC for the EA grant and \$3,200 for the Pell Grant. Refer to the methods section in the text for more details on the estimation. ***p<.01, **p<.05, +p<.1

References

G. Imbens and K. Kalyanaraman. Optimal bandwidth choice for the regression discontinuity estimator. *The Review of Economic Studies*, 2011.

D. Lee and T. Lemieux. Regression discontinuity designs in economics. *Journal of Economic Eiterature*, 48(2): 281–355, 2010.