# ECON 8823 Research Paper:
# Social Security Student Benefit Program's Impact on College Attendance

Haydar Evren[1]

[1]Department of Economics, Boston College

**Abstract**

The research question is whether schooling aid increases college attendance or it just financially supports students who would have gone to college regardless of aid. Examining the effect of schooling aid on college attendance is challenging because treated and untreated students can differ in unobservable characteristics correlated with potential outcomes, even controlling for differences in observed characteristics. In our case, college aid is correlated with many characteristics, treated and untreated may not be directly comparable, even after adjusting for observed characteristics. However, we can use a difference-in-differences estimator for identification strategy. To do that, an exogenous policy change is required. An exogenous change in schooling aid policy can help us in identification. The Social Security Administration has provided benefits to the children of deceased, disabled, and retired Social Security beneficiaries until those children are 18. Between 1965 and 1981, payments were extended to the age of 22 if the child stayed enrolled full time in school. This program was eliminated in 1981. Using difference-in-differences analysis, we find that the availability of $1000 of grant (normalized to $2856) increases college attendance by 0.167 years and the probability of attending by 3.8%.

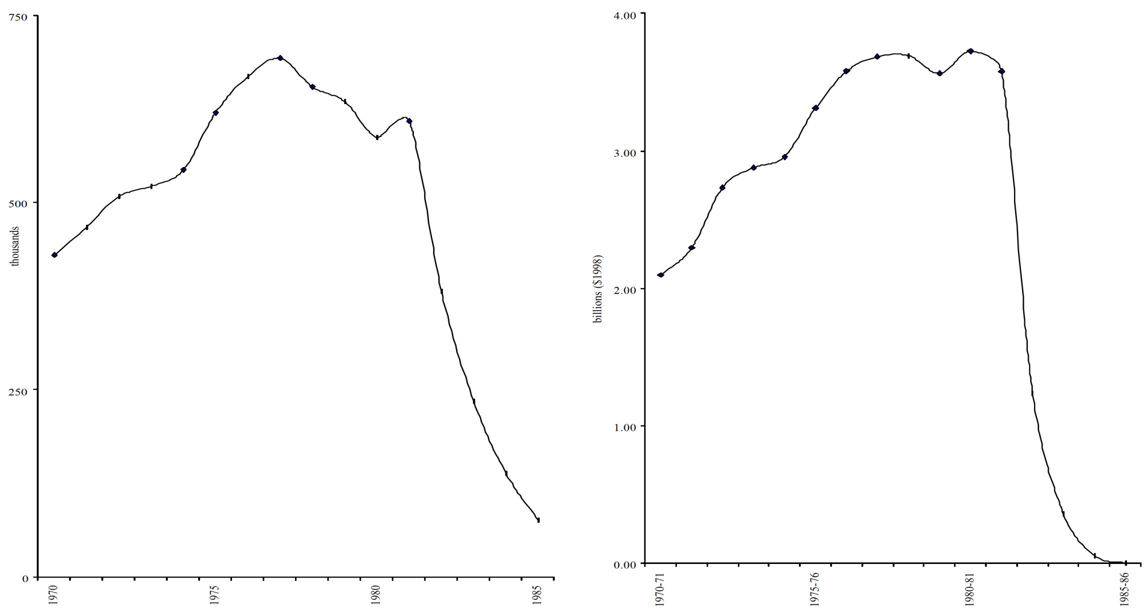
## 1 Introduction

The U.S. spends billions of dollars a year on subsidies for higher education. The bulk of the spending goes to student aid, with the balance going to grants for educational institutions. An increase in student aid is expected to increase higher education enrollment. However, there is little evidence that aid serves its policy goal of increasing college attendance. The billions spent annually on aid may subsidize students who would have gone to college irrespective of aid.

Examining the effect of schooling aid on college attendance is challenging. The straightforward approach is to regress a person's college attendance on the amount of funding she benefits and interprets the coefficient on aid as its casual effect. However, schooling aid is correlated with many attributes that have an impact on educational attainment, and omitting these characteristics from the regression produces a biased estimate. The bias can be eliminated by controlling for observable characteristics that are correlated with both aid and education. However, if we do not add any unobservable characteristics correlated with aid and education, we will end up missing the causal effect of financial aid on college attendance.

To determine the effect of aid, we need an intervention in aid policy that is exogenous to unobservable characteristics that affect educational attendance. A discrete shift in aid policy that affects some students but not others is one such source of exogenous variation. This paper analyzed the effect on college attendance and completed schooling of a significant policy change in federal financial aid policy that happened in 1981.

In the 1965 Social Security Amendments, the definition of a "child" was broadened. In addition to presuming that a child under age 18 was dependent on its parents, the Social Security program began to recognize the reality that children who are full-time students after age 18 are often still, in fact, dependent on their parents for their support. Consequently, the existing Child's Benefit was extended in its duration to include children of the Social Security beneficiary who were full-time students and under the age of 22. The age of 22 was selected because this would be the usual time for a student to complete a four-year college education.




(a) Number of Beneficiaries

(b) Benefits Paid to Beneficiaries

Figure 1: Social Security Administration Collage Student Benefits

Although Social Security student benefits were, in fact, Child's Benefits, they quickly came to be viewed as a form of "student aid," paid, in effect, to help students pursue their education. This popular conception was difficult to displace by any more precise understanding of the social insurance principles underlying the benefit. In any case, the benefits were quite popular. In the peak year of 1977, almost 900,000 students were receiving this type of benefit. In the peak pay-out year of 1981, nearly $2.4 billion was paid in the form of student benefits.[1]

The programs elimination provides an opportunity to measure the effects of financial aid. We find that the removal of the Social Security Student Benefit Program reduced by 27% of the probability that the affected group would complete any years of college. Completed education was cut by 0.59 years. We find that an offer of $1,000 grant increases attainment in education by about 0.16 years and the probability of attending a college by 3.8%.

## 2 Literature Review

We analyze the effect of aid on the probability of a person attending and completing college. The theory that models this relation is simple.

[1]Social Security History. https://www.ssa.gov/history/studentbenefit.html

The classic human capital model, Becker (1993):

$$\max_{E} \int_{t=0}^{E} -Ce^{-rt}dt + \int_{t=E}^{\infty} f(E)e^{-rt}dt$$

where $E$ represents education level, $f(E)$ represents earnings associated with education attained, $r$ represents cost of borrowing and $C$ represents cost of schooling.
Obtaining the first order condition results in:

$$f(E) + C = \frac{f'(E)}{r} \tag{1}$$

And if we introduce aid, we obtain the following equation:

$$f(E) + (C - AID) = \frac{f'(E)}{r} \tag{2}$$

Note that under this simple model of the human capital model, the introduction of aid shifts the optimal equilibrium to a sub-optimal one. However, there are many cases where the individual optimal level of education and social optimal education level might diverge.

The first one is liquidity constraint. If an individual has a binding budget constraint and can not borrow freely, then his education investment is below the socially optimal level. This problem can be solved by introducing student loans and grants for students who have binding borrowing constraints. The second one is student uncertainty about future benefits and costs. Even if a risk-averse individual does not have a liquidity constraint, his investment choice in education will be sub-optimal. In that case, grants and scholarships have a welfare-increasing effect.

The second one is student uncertainty about the costs and benefits of post-secondary education. For example, the return to schooling may change over time due to both shocks to the market for skilled labor and idiosyncratic shocks to a particular degree. Also, a student may be uncertain about his ability to complete college. Even in the presence of student loans and grants, risk-averse individuals will invest in a level of education lower than the socially optimum level. If the government is less risk-averse than individuals, a grant toward schooling costs will have positive welfare effects. Finally, since education can produce externalities, the socially optimum and individual optimum levels always diverge.

The first-order condition (2) implies that heterogeneity in education level choice comes from the level of aid an individual is offered. Therefore, we have this reduced form given below:

$$E_i = \alpha_0 + \beta_0 AID_i + \varepsilon_i \tag{3}$$

The equation implies if schooling aid is uncorrelated with other factors that affects schooling, then $\beta_0$ can be explained as the effect of a dollar increase of aid eligibility on education attainment.

However, it is known that many characteristics that affect educational attainment are also correlated to an individual's eligibility. The fraction of poor students attending college is low even if the aids offered to them are relatively large. In that case, the estimation for the reduced form will give a downward biased estimator. Successful students are also provided generous scholarships. In that case, the estimator for the effect of aid eligibility on education attainment will be biased upward. Therefore, we cannot know how well the estimator in (3) states the casual effect.

To overcome this identification problem, researchers add many covariates.

$$E_i = \alpha + \beta AID_i + \gamma X_i + \nu_i \tag{4}$$

where $X_i$ is a set of variables correlated with both aid and attendance. Many papers are using that kind of analysis. A well know one is College Choice in America by Manski and Wise(1983). Using the data from the National Longitudinal Study, the authors present a set of interrelated analyses of student and institutional behavior, each focused on a particular aspect of the process of choosing and being chosen by a college. Among their interesting findings, schooling aid related one is that Federal scholarship aid has had only a small effect on enrollments at four-year colleges but a much stronger effect on attendance at two-year colleges. They find that $1,000 in Pell grant eligibility increases college attendance by 3.8%.[2]

However, due to information unavailability, we do not observe all covariates. Examples of such covariates are the performance in high school, access to information about the college, level of available financial assets, and the number of dependent siblings or siblings in college. Such characteristics can be modeled as an individual or group-specific error term that is correlated with aid. By taking differences within groups, we can eliminate the source of bias.

Angrist(1993) examines veterans' benefits that subsidize education and training by using variation over time in benefits. Using data from the 1987 Survey of Veterans, these benefits are estimated to increase schooling by 1.4 years.

Manski and Wise (1983), Hansen (1983), Kane (1996), Kane (2010), Ehrenberg and Sherman (1984), Leslie and Brinkman (1987) are some of the early studies examining the effect of the Pell Grant on educational attainment. They failed to find any significant positive enrollment effect; Pell Grants have not improved enrollment rates among low-income students and minorities, but they have likely affected which colleges students choose to attend. [3]

# 3 Social Security Student Benefits

Social Security Act of 1935 contained no provisions for the payment of any type of dependents' benefits. However, even before monthly payments began, the law was significantly changed in 1939 to transform the program into a family-benefits social insurance system. In addition to benefits for the survivors of deceased workers, the program was broadened to include dependents' benefits paid to the spouse or minor children of the retired worker. After disability benefits were added to the program in 1956, these same types of dependents' benefits eventually became available to the families of disabled workers as well.

In the 1965 Social Security Amendments, the definition of a "child" was broadened. In addition to presuming that a child under age 18 was dependent on its parents, the Social Security program began to recognize the reality that children who are full-time students after age 18 are often still, in fact, dependent on their parents for their support. Consequently, the existing Child's Benefit was extended in its duration to include children of the Social Security beneficiary who were full-time students and under the age of 22. The age of 22 was selected because this would be the usual time for a student to complete a four-year college education.

The benefits were quite popular. In the peak year of 1977, almost 900,000 students were receiving this type of benefit. In the peak pay-out year of 1981, almost $2.4 billion was paid in the form of student benefits. Although these benefits were popular with the students and their parents, there were at least three problems with student benefits.

The first problem was the relatively large volume of overpayments experienced in the program. The second problem with student benefits was their cost, in a period when the Social Security program was facing budget pressures.

When the Reagan Administration took office in early 1981, it offered a comprehensive budget and tax proposal designed to achieve its economic objectives. Under the Omnibus Budget Reconciliation Act of 1981, student benefits for post-secondary and elementary or secondary students older than 18 were phased-out and

[2] Manski CF, Wise DA. College choice in America. Harvard University Press; 1983.
[3] Bettinger, E., 2010. Need-based aid and student outcomes: The effects of the Ohio College Opportunity Grant.

finally eliminated by April 1985.[4]

# 4 Methodology and Data

## 4.1 Methodology

The difference-in-differences methodology helps us find the effect of eligibility for Social Security student benefits on college attendance and the probability of attending college. Difference-in-differences is a design that uses longitudinal data from treatment and control groups to obtain a counterfactual to estimate a causal effect. It estimates the effect of a specific policy intervention or treatment by comparing the changes in outcomes over time between a population that is enrolled in a program (the treatment group) and a population that is not (the control group). It relies on a less strict exchangeability assumption. It requires that in the absence of treatment, the unobserved differences between treatment and control groups are the same over time. [5]

Difference-in-difference estimation also requires the parallel trend assumption. It ensures the internal validity of the model. It requires that the difference between the treatment and control group is constant over time in the absence of treatment. Violation of parallel trend assumption will lead to biased estimation of the causal effect. This assumption is tested at the end of the results.

A student is eligible for Social Security Student Benefits for one of three reasons: the death, disability, or retirement of a parent. In the empirical analysis, the treatment group is restricted to those who potentially eligible for Social Security benefits as a result of the death of a parent because the latter two can cause an endogeneity problem: Change in disability benefits in 1980 affected the income level of the disabled-worker families. A parent's decision to enter (or exit) the disability or retirement rolls might correlate with the student benefit's availability. Since the death of a parent is exogenous, it cannot cause an endogeneity problem. We focus on fathers because around 90 percent of students were eligible for benefits through their fathers.

The key estimating equation is the following:

$$E_i = \alpha + \beta(FatherDeceased_i \times Before_i) + \delta FatherDeceased_i + \theta Before_i + v_i \tag{5}$$

where $E_i$ is a measure of educational attainment. $Before$ is a binary variable that is set to one if a student is in the cohort that graduated from high school before college benefits were eliminated. $FatherDeceased$ is a binary variable set to one for those who, due to the death of their father, were potentially eligible for benefits.

The reduced-form effect of Social Security student benefits is captured by $\beta$. The specification controls for changes over time in average college attendance rates and average differences in the college attendance of those with a deceased father and those with a living father. The fundamental identifying assumption is that any relative change in the attendance of the children of deceased fathers is considered to be resulting from eliminating the student benefits. Note that beta captures the effect on schooling decisions of aid eligibility. It is the parameter of interest to predict the impact of changing aid policy.

## 4.2 Data

National Longitudinal Survey of Youth (NLSY) is a survey of young men and women spanning the years 1979 to 1996.[6] The necessary variables for years 1979-1996 from the Department of Labor website are acquired.

[4] Social Security Administration, https://www.ssa.gov/history/studentbenefit.html

[5] Difference-in-Difference Estimation — Columbia University

[6] National Longitudinal Surveys, https://www.nlsinfo.org/content/cohorts/nlsy79

There are 12,686 respondents in the data, out of which 1280 military observations are dropped. Only the respondents who were high school seniors in 1979, 1980, 1981, 1982, and 1983 are kept, which reduces the number of observations to 4537. Also, 355 observations get dropped out for the respondents who had left the survey before 1988 when the questions regarding deceased father were asked, bringing the number of observations used in the analysis to 4182. The variables used are described in Appendix A.1.

# 5 Results

## 5.1 Summary Statistics

The summary statistics are given in Table 1 below.

Table 1: Summary Statistics

| | High School Seniors 1979-1981 | | High School Seniors 1982-1983 | | |
|---|---|---|---|---|---|
| | Father Not Deceased | Father Deceased | Father Not Deceased | Father Deceased | Differences-in-difference |
| Household Income | 25480.10<br>( 13400.90) | 15421.51<br>( 9258.94) | 27729.73<br>( 14758.58) | 15803.68<br>( 11488.02) | 1867.46<br>( 2344.06) |
| Black | 0.14<br>( 0.34) | 0.23<br>( 0.42) | 0.15<br>( 0.35) | 0.25<br>( 0.43) | -0.01<br>( 0.06) |
| Hispanic | 0.05<br>( 0.22) | 0.07<br>( 0.25) | 0.06<br>( 0.24) | 0.05<br>( 0.23) | 0.02<br>( 0.03) |
| Father Attended College | 0.33<br>( 0.47) | 0.18<br>( 0.39) | 0.30<br>( 0.46) | 0.13<br>( 0.34) | 0.02<br>( 0.07) |
| Mother Attended College | 0.24<br>( 0.43) | 0.13<br>( 0.33) | 0.21<br>( 0.40) | 0.14<br>( 0.35) | -0.04<br>( 0.08) |
| Single-Parent Household | 0.20<br>( 0.40) | 0.85<br>( 0.36) | 0.24<br>( 0.43) | 0.90<br>( 0.30) | -0.01<br>( 0.06) |
| Family Size | 4.71<br>( 1.66) | 4.34<br>( 2.11) | 4.61<br>( 1.71) | 3.92<br>( 2.13) | 0.31<br>( 0.33) |
| Age in 1988 | 25.51<br>( 1.05) | 25.57<br>( 1.15) | 23.52<br>( 0.79) | 23.42<br>( 0.69) | 0.15<br>( 0.16) |
| Female | 0.48<br>( 0.50) | 0.50<br>( 0.50) | 0.47<br>( 0.50) | 0.49<br>( 0.50) | -0.00<br>( 0.09) |
| Attend College by 23 | 0.55<br>( 0.50) | 0.59<br>( 0.49) | 0.54<br>( 0.50) | 0.37<br>( 0.49) | 0.21**<br>( 0.09) |
| Complete any College by 23 | 0.48<br>( 0.50) | 0.56<br>( 0.50) | 0.46<br>( 0.50) | 0.31<br>( 0.47) | 0.23**<br>( 0.09) |
| Years of School at 23 | 13.39<br>( 1.82) | 13.34<br>( 1.70) | 13.32<br>( 1.90) | 12.70<br>( 1.51) | 0.58**<br>( 0.29) |
| Number of Observations | 2883 | 145 | 1094 | 60 | 4182 |

Notes: Standard errors are in parentheses. The standard errors in the differences-in-differences column are adjusted for clustering at the household level. ***, **, and, * indicate statistical significance at the 1, 5, and, 10 percent level respectively.

Students with deceased fathers grow up in relatively low-income families and are more likely to live in single-parent households. Students with deceased fathers are more likely to be black, due to the higher mortality rate of prime-age black men. For the cohort of students who were high school seniors in 1979, 80 and 81, those with deceased father were more likely to attend college: 59% had participated in college, while only 55% of seniors with living fathers had done so. For senior students in 1982 and 1983, the pattern is changed: only 37% of seniors whose fathers had died by the time they were 18 went to college, whereas

54% of their classmates attended. So Table 1 provides and evidence that the additivity assumption of difference-in-differences holds for this analysis.

## 5.2 Effect on Probability of Attending College

We use ordinary least squares (OLS) to estimate the key equation (5). Standard errors are adjusted for heteroskedasticity and multiple observations within households. Failure to do so can lead to some misleading outcomes: small standard errors, narrow confidence intervals, large t-statistics and low p-values.

Table 2: OLS, Effect of Eligibility for Student Benefits on Probability of Attending College by 23

| | Differences-in-Differences (1) | Add Covariates (2) |
|---|---|---|
| Deceased Father × Before | 0.208** | 0.250** |
| | (0.09) | (0.11) |
| Deceased Father | -0.167** | 0.982 |
| | (0.08) | (0.74) |
| Before | 0.009 | -0.383 |
| | (0.02) | (0.42) |
| Household Income | | 0.000** |
| | | (0.00) |
| Black | | 0.216*** |
| | | (0.04) |
| Hispanic | | 0.166*** |
| | | (0.04) |
| Father Attended College | | 0.188*** |
| | | (0.04) |
| Mother Attended College | | 0.104** |
| | | (0.04) |
| Single-Parent Household | | -0.095*** |
| | | (0.04) |
| Family Size | | -0.017** |
| | | (0.01) |
| Age in 1988 | | -0.058*** |
| | | (0.01) |
| All Covariates × Before | | Yes |
| All Covariates × Deceased | | Yes |
| Region Dummies | | Yes |
| R Squared | 0.002 | 0.296 |
| No. of Observations | 4182 | 4119 |

Notes: Regressions weighted by 1988 sample weights. Standard errors adjusted for heteroskedasticity and multiple observations within households. ***, **, and, * indicate statistical significance at the 1, 5, and, 10 percent level respectively.

In the second column, we include other covariates and run a similar specification. We add dummies for the region where the variable on a region is taken from the survey in 1979. Additionally, to account for the bias caused by heterogeneity across time and eligibility status, the specification includes interactions of covariates with the 'before' dummy and the 'father deceased' dummy. The estimated effect of aid eligibility on attendance barely changes with the addition of covariates: it is 25 percentage points, with a standard error of 11 percentage points, and is still significant at 5% level. Also, the explanatory power of the regression rises dramatically from 0.002 to 0.296.

One thing to note is that the coefficient on the 'father deceased' dummy flips sign and becomes insignificant on adding covariates. This could be possible if, for instance, people living in one region had higher mortality and weaker educational outcomes. In the absence of control for the region, it would lead to a negative sign for the coefficient even when it is possible that the real underlying relation was positive. A positive relationship could hold if, for instance, the students whose father passed away felt more motivated than others and were more likely to go to college.

The puzzling thing about Table 2 is that most of the other covariates are significant at a 1% level of significance. We do not know what to conclude from it apart from the fact that these covariates are possibly the key determinants of college attendance, or we have made some mistakes while working with the data. The positive sign on the coefficient for Black and Hispanics is shocking. However, the sign on the coefficient of the other covariates does make sense. Higher family income and having parents who attended college increase the probability of attending college while living in a single-parent household and having a larger family leads to a lower probability of going to college.

## 5.3 Effect on Other Schooling Outcomes

Table 3 is the estimation of the effect of aid eligibility on completed schooling in addition to college entry. The estimates are based on the fully controlled specification of Table 2.

Table 3: OLS, Effect of Eligibility for Student Benefits on Schooling Outcomes

| | Attended College Full-Time (1) | Completed Any Years of College (2) | Years of Schooling (3) |
|---|---|---|---|
| By Age 23 | 0.25** | 0.27** | 0.59 |
| | (0.11) | (0.11) | (0.37) |
| By Age 28 | | | |
| Exclude Attriters | 0.28** | 0.29** | 0.84* |
| | (0.11) | (0.12) | (0.45) |
| Assign Last Values | 0.26** | 0.26** | 0.75* |
| | (0.11) | (0.11) | (0.43) |

Notes: Coefficients are those on deceased father × before in regressions in which the outcomes are those indicated in the columns. The regression specification is that of column (2) in Table 2. ***, **, and, * indicate statistical significance at the 1, 5, and, 10 percent level respectively.

Eligibility for student benefits appears to increase the probability of completing at least a year of college by 27 percentage points and is significant at 5% level of significance. The coefficient on years of schooling is not significant.

It is possible that the effect of aid eligibility could be leading students to go to college earlier rather than increasing their school investments. Table 3 also reports the results of estimation from the same specification but now with schooling decisions at the age of 28. The attrition between age 23 and 28 can bias the results. Around 8% of respondents in the final dataset exited the sample by age 28. We did not worry about respondents leaving the sample before they were 23 because our sample includes only respondents who were there in 1988 when all of them were between ages 23-31. We use two approaches to dealing with attrition: Dropping the attriters and assigning them their last observed value of the dependent variable. In both cases, estimates at age 28 are not lower than those at age 23. This suggests that aid eligibility did not simply speed up investment in schooling but also raised its optimal level.

## 5.4 Estimation using Probit and Logit

The differences-in-differences estimate can change the sign if a nonlinear transformation, such as log, is applied to the dependent variable. The estimates here are not vulnerable to this functional form sensitivity as the children of deceased fathers were more likely than their counterparts to attend college before the policy change but less likely after. Under these conditions, linear and non-linear estimates will be of the same sign.

Table 4: Probit and Logit, Effect of Eligibility for Student Benefits on Probability of Attending College by 23

| | Coefficients<br>(1) | Marginal Effects<br>(2) |
|---|---|---|
| Probit Estimates | 0.53** | 0.21 |
| | ( 0.24) | |
| Logit Estimates | 0.85** | 0.21 |
| | ( 0.39) | |

Notes: Regressions weighted by 1988 sample weights. Standard errors adjusted for heteroskedasticity and multiple observations within households. ***, **, and, * indicate statistical significance at the 1, 5, and, 10 percent level respectively.

Table 4 reports estimates from probit and logit estimation of the effect of aid eligibility on the probability of attending college. The specification used is same as Column (1) of Table 2. The coefficients from both the regressions tell the same story as the linear probability model and are significant at 5% level of significance. The predicted marginal effect of the coefficient of interest is presented in Column (2). The OLS estimates and predicted probit and logit marginal effects are similar at 21 percentage points.

Note that the marginal effects calculated in Column (2) of Table 4 are the following

$$MarginalEffect = F(\alpha + \beta + \delta + \theta) - F(\alpha + \delta + \theta)$$

where $F$ is normal and logistic distribution for the probit and logit estimation, respectively. Note that is the actual difference-in-difference estimator for non-linear estimations of these sort. [7]

## 5.5 Falsification Test

A critical assumption in differences-in-differences approach is the parallel trend. By showing that the difference-in-difference for all control variables is insignificant, the two group - seniors with deceased fathers and seniors with living fathers- would have evolved similarly in the absence of policy. (We could relate this argument to a Regression Discontinuity design - that is, within a small neighborhood around the cutoff, assignment to treatment is random, and thus discontinuity of outcome at the cutoff gives us the treatment effect. This is valid only if we observe the discontinuity just in the outcome of interest but not in background characteristics.) Essentially this argument claims that nothing else in the same time period could have induced differential changes in outcomes. However, this argument is not persuasive because we never know

[7] Puhani, Patrick A. The Treatment Effect, the Cross Difference, and the Interaction Term in Nonlinear "Difference-in-Differences" Models. IZA Discussion Paper Series

what would have happened to the schooling outcomes of seniors with deceased fathers had the aid policy never been implemented.

Another way to check the assumption is to check for pre-existing trends. However, the NLSY survey started in 1979, while the aid program had been in place since 1965, so we cannot test for trends in the pre period. What we do have is data even after 1983, so we can conduct a placebo test for the post period. With NLSY data, we can compose an additional cohort of seniors between 1984 and 1990. We still compare seniors with deceased father versus those with living father, but use the 82-83 cohort as before and 84-98 cohort as after. Since the aid benefits were withdrawn in 1981, from 1982 onwards, the policy no longer existed and a fictitious treatment to seniors with deceased father in the before cohort should have no effects on schooling outcomes.

Table 5 presents the results from a simple falsification test. We use the shufflevar command in Stata and shuffle the before variable so that now youth in both father deceased and father not deceased group can randomly be high school seniors in any year from 1979-1983. The specification used for estimation is the same as in Table (2). The coefficients in Table (5) have the wrong sign and are not significant. This once again confirms that having a deceased father increased a persons probability of attending college only if they were a high school senior before the policy change.

Table 5: Falsification Test

| | Differences-in-Differences (1) | Add Covariates (2) |
|---|---|---|
| Deceased Father × Before | -0.137 | -0.123 |
| | (0.10) | (0.09) |
| Deceased Father | 0.067 | -0.467 |
| | (0.08) | (0.71) |
| Before | 0.012 | 0.242 |
| | (0.02) | (0.34) |
| R Squared | 0.001 | 0.293 |
| No. of Observations | 4182 | 4119 |

Notes: ***, **, and, * indicate statistical significance at the 1, 5, and, 10 percent level respectively.

## 6 Conclusion

Using difference-in-differences estimator, we find that the availability of $1000 of grant (normalized to $2856) increases college attendance by 0.16 years and the probability of attending by 3.8%. After elimination of the aid completed education was reduced by 0.59 year. Moreover, we also find that aid eligibility does not simply speed up investment in schooling but also raises its optimal level.

## References

J. D. Angrist. The effect of veterans benefits on education and earnings. *ILR Review*, 46(4):637–652, 1993.

G. S. Becker. Human capital revisited. In *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education (3rd Edition)*, pages 15–28. The university of Chicago press, 1994.

E. Bettinger. Need-based aid and student outcomes: The effects of the ohio college opportunity grant. *Unpublished manuscript, Stanford University School of Education. Retrieved from http://www. sesp. northwestern. edu/docs/need-basedaid-why. pdf*, 2010.

R. Bifulco, J. M. Fletcher, and S. L. Ross. The effect of classmate characteristics on individual outcomes: Evidence from the add health. 2008.

J. T. Denning, B. M. Marx, and L. J. Turner. Propelled: The effects of grants on graduation, earnings, and welfare. Technical report, National Bureau of Economic Research, 2017.

S. M. Dynarski. Does aid matter? measuring the effect of student aid on college attendance and completion. *American Economic Review*, 93(1):279–288, 2003.

R. G. Ehrenberg and D. R. Sherman. Optimal financial aid policies for a selective university. *Journal of Human Resources*, pages 202–230, 1984.

W. L. Hansen. Impact of student financial aid on access. *Proceedings of the Academy of Political Science*, 35(2):84–96, 1983.

T. J. Kane. Lessons from the largest school voucher program ever: Two decades of experience with pell grants. *Who chooses*, pages 137–150, 1996.

T. J. Kane. *The price of admission: Rethinking how Americans pay for college*. Brookings Institution Press, 2010.

L. L. Leslie and P. T. Brinkman. Student price response in higher education: The student demand studies. *The Journal of Higher Education*, 58(2):181–204, 1987.

V. Y. T. Liu. The impact of year-round pell grants on academic and employment outcomes for community college students. 2017.

C. F. Manski and D. A. Wise. *College choice in America*. Harvard University Press, 1983.

S. K. Riegg. Causal inference and omitted variable bias in financial aid research: Assessing solutions. *The Review of Higher Education*, 31(3):329–354, 2008.

N. S. Seftor and S. E. Turner. Back to school: Federal student aid policy and adult college enrollment. *Journal of Human resources*, pages 336–352, 2002.

# A Appendix

## A.1 Data

The variables used in the analysis are described below:

*Before*: Download the variable (q3_1c_1_year) that tells what grade the respondent is currently enrolled in. Then generate a variable for each respondent which takes the value of the year whenever the respondent's currently enrolled grade is 12. This is the year in which the youth is a high school senior. Code 'before' as one if the high school year is 1979, 1980 or 1981 and code it to 0 if the high school year is 1982 or 1983.
*Father Deceased*: Download variable (cres_2b_i_1988) which tells times and the parent the respondent stopped living with and variable (cres_2c_i_1988) which tells the reason as to why the respondent stopped living with the parent at that time. If the first variable is equal to the father and the reason is death, then we know if the respondent's father died. We do indeed find that in the data approximately 5% respondents lost their father.
*Sampling Weights*: Obtain from the 1988 sample
*Household Income*: Download the variable (tnfi_trunc_year) which gives the total net family income. We also replace the missing values with the cohort-specific means.
*Black and Hispanic*: Generate from the sample race variable from the 1979 survey.
*Father Attended College and Mother Attended College*: Obtained from highest grade completed by father and mother from the 1979 survey.
*Single Parent Household*: Download the variable (hhi_final_relcode_i_year) which tells the relationship of each individual in the household to the respondent. If any of the members of the household is respondent's mother/foster-mother/step-mother, then we know if the respondent had a mother in any given year. Similarly, we can find out if the respondent had a father in any given year. From this, we can code if the respondent lived in a single parent household when he/she was a high school senior.
*Family Size*: Download the variable (famsize_year) and use it for the year in which the respondent is a high school senior.
*Age in 1988*: Age in 1988 = Age in 1979 + 9
*Female*: Obtained from sample sex variable in 1979
*Attend College by 23*: Download the variable (enrol_year) and code 'Attend college by 23' as 1 if the respondent has been enrolled in college at any time before the age of 23.
*Complete Any College by 23*: Download the variable (q3_3_year) that tells the highest grade completed in any year. If highest grade completed is greater than or equal to 13 in any year when the respondent is less than or equal to twenty-three years old then code this variable as 1.
*Years of school at age 23*: Download the variable (q3_3_year) for the highest grade completed in the year and take the maximum value found from any year when the respondent is less than or equal to twenty-three years old.