Does Higher Education Cause Political Participation?: Evidence From a Regression Discontinuity Design*

Alex Solis†

October 26, 2015

Abstract

The relationship between education and political participation is one of the most debated in economics. However, only few papers have explored the causal link with contradictory findings. I use the eligibility criteria for higher education loans in Chile that produce exogenous variation on enrollment to test causality on political participation using an RDD. Using administrative individual data from the universe of voters complemented with survey data from a representative sample, I find that higher education does not cause political participation This is confirmed with survey data for several other forms of political participation and sophistication. However, higher education causes overreporting.

JEL Codes: D7, I21, I25.

Keywords: political participation, college, higher education, voting registration, overreporting.

*I would like to thank David Card, David Figlio, Frederico Finan, Ernesto Dal Bó, Eva Mörk, Enrico Moretti, Mattias Nordin, Elizabeth Sadoulet, Emmanuel Saez, and seminar participants at Institute for Evaluation of Labor Market and Education Policy, IFAU; LACEA 2012 Annual Meeting; PUC Chile; Stockholm School of Economics; The Swedish Institute for Social Research, SOFI; UC Berkeley ARE Development Workshop; University of Gothenburg; University of Naples Federico II, who provided useful comments and insights. I would like to thank Manuel Matta, Francisco Pino, Gonzalo Sanhueza, Teresa Veloso, and Humberto Vergara for providing the data. I gratefully acknowledge financial support from the Center for Equitable Growth at the University of California, Berkeley. All errors are my own.

†Email: alex.solis@nek.uu.se. Department of Economics, Uppsala University P.O. Box 513 75120 Uppsala, Sweden and Uppsala Center of Labor Studies.

1 Introduction

The relationship between education and political participation is one of the most debated in economics: in the growth literature, human capital accumulation is seen as a necessary condition for democracy enabling prosperity and growth (e.g., Barro (1999), Glaeser et al. (2007); and Acemoglu et al. (2005)); in the economics of education, political participation is considered one of the most important externalities (Friedman (1962), Hanushek (2002), Moretti (2004) and Oreopoulos and Salvanes (2011)); and in political economics and political science, it is considered one of the most important relations (Lipset (1959) and Campante and Chor (2012)), and sometimes, it is considered one of the major contributions of political science to the general body of knowledge (Berinsky and Lenz (2010)).

The strong and positive relationship it is one of the most well documented empirical regularities (Campbell et al. (1960), Wolfinger and Rosenstone (1980), Nie et al. (1996), among others), however, causal evidence, at the individual level, is scarce and controversial. For instance, using US data, Dee (2004), Milligan et al. (2004) and Sondheimer and Green (2010) find positive and significant effects,[1] while Tenn (2007), Kam and Palmer (2008), and Berinsky and Lenz (2010) find zeros.[2,3] The estimation of this relationship suffers from at least two forms of endogeneity. First, unobservable variables may induce individuals to self-select into higher education and into political activities simultaneously (e.g., personality traits, cognitive ability, family background, etc.) leading to biased estimates.[4] Secondly, it may suffer from reverse causality, i.e., more politically active communities have the tools to demand for policies in favor of more and better education. The controversy in the results has centered on validity of the methods implemented. For example, Tenn (2007), Mazumder (2008) and Stephens and Yang (2014) have challenged the validity of the results using compulsory schooling laws as instrumental variable, studying a more

[1] Dee (2004) and Milligan et al. (2004) use instrumental variables (exposure to labor laws and compulsory schooling respectively). Nevertheless, Milligan et al. (2004) apply the same instruments to the United Kingdom finding a non-statistically significant effect. Sondheimer and Green (2010) analyze three experiments (Perry Preschool, Tennessee STAR and "I Have a Dream" scholarship) finding a positive and significant relationship in only one.

[2] Tenn (2007) exploits differences in years of education for same age individuals, due to requirements to enter the educational cycle; Kam and Palmer (2008) use a propensity score matching; and Berinsky and Lenz (2010) use the Vietnam draft as instrument for education.

[3] Additionally, using data for other countries and compulsory schooling laws as instrumental variables, Borgonovi et al. (2010), Pelkonen (2010), Siedler (2010) and Chevalier and Doyle (2012) find no effects. Wantchekon et al. (2015) shows that schools have an effect on political participation in Benin.

[4] Kam and Palmer (2008) argues that education acts as a proxy.

general specification that includes region-specific time trends.[5]

In this paper, I deal with endogeneity using a regression discontinuity design (RDD). Students in Chile became eligible for higher education loans in any type of institution when they score more than a given threshold in the national college admission test (Prueba de Selección Universitaria, PSU hereafter), while students below the cutoff are only eligible for loans in vocational and technical institutions. At the threshold, this rule causes an important change in higher education enrollment and, particularly, in college that is as good as randomly assigned (Lee (2008)). I use this exogenous variation in enrollment to estimate the causal effect of higher education (and college) on political participation.

I find that, despite the positive and strong correlation between higher education and voting registration in the sample, the causal relationship is precisely zero. The results uses individual level administrative data in three dimensions: 1) the universe of participants in the college admission process,[6] 2) the universe of enrolled students in all higher education institutions across the country and, 3) the electorate registry. Results are robust to different specifications, the inclusion of a rich set of covariates. Moreover, the relationship is zero in all analyzed sub-samples (by income level, gender or family background characteristics).

Additionally, I linked the administrative data with information from a survey from a representative sample around the loan cutoff, to explore the effects of higher education on other forms of participation that lack of administrative data, e.g., other political activities, demonstrations, attitudes towards democracy, etc. I find that the relationship is zero for all the explored forms of participation and for different forms of political sophistication.

The analysis of higher education is relevant for the discussion of the relationship, because can be linked to the two main mechanisms argued in the literature through which education can affect political participation: the cognitive ability hypothesis (or civic education hypothesis) and the social network hypothesis. The former states that education prepares individuals to understand the political process, giving individuals the necessary skills and knowledge to get involved and to reduce the cost of participation (Wolfinger and Rosenstone (1980)). College life is rich in activities that give individuals the opportunity to understand the importance of politics and

[5]See also Acemoglu et al. (2005) vs Bobba and Coviello (2007) and Castelló-Climent (2008)

[6]Approximately, 80% of all high school graduates in the country take the PSU test.

highlight the values of democracy, allowing the emergence of tastes for participation (Galston (2001)). The social network hypothesis argues that the influence of education is relative, and higher education, and college in particular, puts individuals closer to political networks and the spheres of power encouraging participation. Also, peers in social networks provide individuals with political information reducing the cost of participation (Nie et al. (1996)). Social networks can also impute social pressure to engage political behavior (Funk (2010)) and may be the target of politicians who try to maximize the outcome of their information campaigns (Hillygus (2005)). This paper contributes to the literature testing these mechanisms.

I investigate the importance of the cognitive ability hypothesis, testing how much individuals know about the political system and key politicians. I find that higher education do not cause more knowledge.[7] Additionally, I test the hypothesis of social networks, showing evidence of a causal relationship between higher education and overreporting, which is consistent with the existence of social pressure due to social norms (e.g., Silver et al. (1986), Bernstein et al. (2001), Milligan et al. (2004) and Karp and Brockington (2005)). I construct a measure of overreporting combining the official registration to vote with the self-reported in the survey. I find that students scoring barely above the loan cutoff are two percentage points more likely to overreport than their counterfactual barely below. 2SLS results suggest that college would increase the probability of overreporting in 10% (16% for enrollment in any higher education).

To the best of my knowledge, the contributions of the paper are three fold. First, in term of the identification strategy, it is the first paper using a regression discontinuity design to explore the causal link between education and political participation. Second, in terms of the data, it is the first paper that links the whole electorate with the universe of students enrolled in higher education at the individual level, in a setting with an arguably credible identification strategy. Moreover, combining the administrative data to a survey allows me to analyze alternative measures of participation. Finally, is the first paper that tests causally the channels through which education affects political participation presented in the literature.

The paper is organized as follow. In Section 2, I describe the electoral system, the institutions in higher education, and the data used in the paper. Section 3 describes the identification strat-

[7] A key element of this test is that knowledge cannot be overreported, which is ensured by questions that use pictures to restrict the ability of finding the right answer in the web.

egy. Section 4 presents the empirical results for registration to vote using administrative data. Section 5, uses survey data to show evidence for other forms of political participation and political attitudes. Also, shows evidence for the two channels that can be analyzed with the data. Finally, Section 7 concludes.

2 Background

2.1 Electoral System

In Chile, eligible voters are entitled to vote in three types of elections: presidential, parliamentary and municipal elections.[8]. Until the presidential election of 2009, voting age individuals needed to voluntarily register to be entitled to vote. After registration, voting was mandatory with fines that were high relative to the median income.[9] In the following election of 2012, the system changed to voluntary vote for voting age individual. Because this change, I use the registry in 2009 as a measure of political participation. This registry indicates that 71% of the voting-age population was registered to vote,[10] with turnout slightly below 90% among registered.[11] To be eligible to register, citizens and permanent residents needed to be at least 18 years old.

2.2 Higher Education in Chile

Higher education in Chile is offered in three types of institutions: vocational institutions (Centros de Formación Técnica), technical institutions (Institutos Profesionales), and universities (which include teaching and research universities).[12,13] In all types, the curriculum is very rigid with predetermined subjects and little space for choices.

[8] All elections are held on Sundays (a holiday) and poll stations are located in schools and stadiums well known by the electorate. Elections for president, half of the senate and the chamber of deputies are held at the same day every four years. Municipality elections to elect mayor and its council, are held one year before (every 4 years).

[9] Fines for not voting were roughly $200 USD of 2009 (median wage among employed individuals was $720 USD), unless the individual is 200 kilometers away from her poll station, sick, or has lost her national id card, in which case the individual needs to hereby certify at a police station to avoid the penalty. Nevertheless, certifying the inability to vote could imply a similar problem in terms of time and transportation, since a few police stations were available to certify the problem.

[10] Using projected population from Census 2002 and actual registration data from electoral commission.

[11] Figure 4 in the Appendix shows that Chile is a country with average turnout very similar to the US, and very similar in terms of population in higher education with European countries like France and Austria.

[12] Throughout the paper I use college and university as synonyms.

[13] In 2007, vocational institutions serve 12% of the total higher education enrollment, technical institutions 20%, universities 68%. (Rolando et al. (2010))

The key differences between these institutions are the type of degree, the graduation time and the tuition costs. Vocational programs prepare students to become assistants and support staff in different specialties. The majority of these programs last for two years and they grant certificates of technical higher education (títulos técnicos de nivel superior).[14] Technical institutions are similar to community or junior colleges in the US. Their programs grant professional titles (títulos profesionales) that are not bachelor degrees. These programs last for three to four years, and are manly focused in specific knowledge to perform blue collar jobs. A university is the only one able to grant a bachelor or academic degrees (licenciatura). Most of university programs last five years, even though the average graduation time is six.[15] The instruction is more holistic and covers an ample range of topics.[16] There is two types of universities: traditional and private, and both use the PSU test to select students. The only reason to make the distinction is that one of the loan programs analyzed in the next section is exclusively for traditional universities.[17]

In terms of political participation, only university students are organized in student unions, with highly political elections, oftentimes, representing the national political parties. Universities also, have been used as battle ground to demand for political changes. Since 2006, universities have been continuously striking, demanding for a change in the educational system at all levels, and other social and political demands.

To enroll in any higher education institution, students need to graduate from high school. For the years considered in the paper, roughly 80% of the population graduated from high school, and among them, 80% take the PSU test.[18] Almost all students who enroll in higher education have taken the PSU in the year they graduate from high school.[19] PSU scores are used for

[14]For example, international cook, educational assistant and massage therapist.

[15]Graduation time in vocational and technical institutions is 3.5 years and 4.5 years respectively.

[16]For example, to become an economist, students enroll in a five year long program called "commercial engineer" that teach management, human resources, labor law, accountability and economics. Students choose an specialization only in the last two years.

[17]Traditional universities are 25 institutions founded before the educational reform of 1981. All the other, founded since 1981, are called private, but by law they are non-for-profit.

[18]According the CASEN 2009 (the main household survey in the country), about 99% of the population finish 8th grade. I use registry data on the universe of eighth graders from the Ministry of Education to track high school completion. I merged that data using national ID numbers to PSU tests from 2009 to 2012 to find out who takes the test.

[19]The PSU test is written once a year, a few weeks after the end of the school year by about 250 thousand students per year.

admission in university programs.[20][21] More importantly, PSU scores are necessary to obtain financial benefits from the ministry of education. All loans and grants provided with public funds determine student eligibility based on PSU score cutoffs (with is only one exception). These cutoffs provide a exogenous variation in higher education that is explained in the next section.

2.2.1 Loans as natural experiment

Higher education is costly in terms of per capita income. On average, vocational, technical and university programs cost annually the equivalent of 1.5, 3 and 6 median monthly wages. As a consequence, students rely heavily in loans and grants given by the government.[22] By far, the most important programs are the State Guaranteed Loan program (Crédito con Aval del Estado, SGL hereafter) and the traditional universities loan program (Crédito Solidario para Universidades Tradicionales, TUL hereafter). They served about 50% of the eligible population.

To have access to any of these loans, students need to complete a socioeconomic verification form (Formulario Único de Acreditación Socioeconómica, FUAS) before taking the PSU test, and be classified by the tax authority in one of the four poorest income quintiles.[23] Students complying with these requirements are called "*pre-selected*" for loans.

Pre-selected students scoring at least 475 points in the PSU (average in language and mathematics) are eligible for both of these loans in any type of institution.[24] Student scoring less than 475 are only eligible for SGL in vocational and technical institutions if their high school GPA is larger than 5.2.[25] As a consequence, scoring above 475 point in the PSU test allows students to enroll in a university with higher probability. This feature of the higher education system in Chile enables a regression discontinuity design that will be used for identification (see more in Section 3).

[20]Both types use the PSU scores to select students. Traditional universities have a centralized system that weight PSU scores and High school GPA. Private universities, accepting students with one of the loans from the government, are mandated by law to select according to PSU scores. The resources from this loan are the main source of financing for private universities.

[21]The subjects tested in the PSU are math, language, science and history. The average in the first two is used for aid and loan eligibility, and it is referred as the "*PSU score*".

[22]There is no other important source of financing for higher education, except for scholarships for very good students given by specific universities. Those scholarships do not affect the behavior of the students used around the loan cutoff.

[23]The tax authority uses official tax records to rank families.

[24]The institution needs to be accredited by the higher education commission, and 93% of them are accredited.

[25]GPA scale goes from 1 to 7.

2.3 DATA

In the first part, I use registration to vote as a measure of political participation. As explained in Section 2.1, given the high penalties for not voting after registering, registration to vote is a good proxy for voter turnout, the main form of political participation analyzed in the literature.

To test the effects of higher education on registration to vote, I combine four layers of administrative data for the whole population. First, the PSU data set, which includes the college admission test scores and a rich set of demographics characteristics.[26] Second, higher education enrollment from the Ministry of Education contains enrollment information for all programs and all institutions for the universe of students that have enrolled in the years considered. Third, the FUAS data set contains information on application to student benefits and, more importantly, contains the income quintile reported by the national tax authority that determine eligibility for these loans and grants. Finally, the information on registration to vote corresponds to the electoral registry from the Electoral Commission (Servicio Electoral de Chile) immediately before the presidential election of December 2009.[27]

In the second part, I combined the previously described administrative data with self-reported information elicited through a web survey, specially designed for this paper, and performed in October of 2012.[28] The invitation to answer the survey was sent to email addresses registered by the students in the PSU registration form.

About ten thousand students responded the survey implying a response rate of about 40%, calculated among emails that were opened. The invitation email included a request for a "read receipt," that allows to determine whether the message was opened.[29,30] Section 5 will show that students who answered the survey constitute a representative sample of the population of interest.

[26]Self-reported family income, parent education, household size, place of residency, which is combined with administrative records such as gender, birth date, graduating high school, high school GPA, graduating year, etc. The data comes from DEMRE, the institution in charge of the application of the test.

[27]All data sets were merged using national ID numbers (RUN).

[28]The survey was performed by the Universidad Católica de Concepción, a traditional university that belongs to the Council of rectors of traditional universities (Consejo de rectores de las universidades chilenas, CRUCH), which sent the emails and merged the responses with the administrative data.

[29]According to this measure, about 25,000 email were actually opened. Not all email platforms allow to send a read receipt back to the sender, thus this number are only for reference.

[30]Only 13% of the invitations were open. Emails were sent in in October 2012 to all pre-selected students with valid emails addresses in the PSU data set (about 200,000). Many reasons can explain this percentage. First, these emails addresses were self-reported by students between 2007 and 2009, thus they may be outdated. Second, the invitation email may have ended in spam folders. The sender email was not previously known by the students, and contained words such as "invitation", "survey", "questionnaire", etc., which raise red flags from email servers. Finally, students may have followed IT recommendations to do not open emails from unknown senders.

Appendix E contains the full set of question.

3 Identification Strategy

I address endogeneity using a natural experiment that produces as good as random variation in higher education enrollment. Students who score above the loan cutoff on the PSU test have access to tuition loans for any institution, while the group of students below the cutoff has access to loans for vocational and technical institutions only. If we assume that each individual's PSU score (the assignment variable) is determined by her characteristics (such as family background or latent ability) and a random factor with a continuous density, then, conditional on being in an ϵ-window around the cutoff, τ, the probability of being above or below is the same, for an ϵ sufficiently small (see Lee (2008)).[31] Thus, receiving access to loans for any institution is as good as random in a neighborhood of the threshold, creating exogenous variation in higher education enrollment. Because the probability of receiving the higher education "treatment" jumps discontinuously at the cutoff (with imperfect compliance), assuming monotonicity allows us to interpret the estimates as the causal effect on the compliers (LATE), i.e., those individuals that enrolled into higher education when they score above the cutoff, but would have not otherwise (see Imbens and Angrist (1994) and Hahn et al. (2001)).

In the next section, I will show that the probability of enrolling into higher education and college jump at the cutoff. Since these changes may have a different interpretation, I will start showing reduced form estimations. Following Lee and Lemieux (2010) I will estimate:

$$PolPart_i = \alpha_3 + \gamma \cdot \mathbb{1}((T_i \geqslant \tau) + f(T_i - \tau) + \zeta_i \tag{1}$$

Where $PolPart_i$ is a indicator variable for political participation. $\mathbb{1}((T_i \geqslant \tau)$ is an indicator whether individual i scores (T_i) more than the loan cutoff (τ). Ideally, we would estimate equation (1) for a very small window around the cutoff, however, in general, there are only few observations available, hence, the estimates suffer from small sample biases. To increase precision, one can use

[31] By the continuity of the density function we have that for all δ>0, there is an ϵ such that $Pr(\epsilon < T - \tau < 0) - Pr(0 \leqslant T - \tau < \epsilon) < \delta$, i.e., the probability of scoring ϵ below the cutoff minus the probability of scoring ϵ above can be made arbitrarily small, which is equivalent to random assignment: The probability of being in the control or treatment groups is the same, conditional on having $T - \tau$ in the interval $(-\epsilon, \epsilon)$.

observations away from the cutoff controlling for the influence of the PSU score in the probability of enrollment using a flexible control function f (a different function at each side of the cutoff). If the control function is wrongly specified, the estimates suffer from the bias of using not comparable individuals. For robustness, I show results where f is a fourth order polynomial using the whole PSU support; and for linear f in a bandwidth around the cutoff, which is calculated minimizing the sum of these two biases in a mean squared error fashion, the Imbens and Kalyanaraman (2012) optimal bandwidth (I&K hereafter).[32] ζ is an exogenous error term.

The parameter effect of interest, γ, is interpreted as the "intention-to-treat" effect. Estimating a zero in the reduced form implies a zero effect in the 2SLS estimation (the 2SLS estimator is just the ratio between the reduced form estimation and the size of the first stage).

To discuss the size of the effects and to interpret the effects of higher education, I show 2SLS estimates. In some cases, I start showing how strong variation in higher education is at the cutoff (the first stage), controlling for the influence of the running variable in a flexible manner as before:

$$Education_i = \alpha_1 + \beta_1 \cdot \mathbb{1}((T_i \geqslant \tau) + g(T_i - \tau) + \nu_i \tag{2}$$

$Education_i$ is a dummy whether individual i enrolled in any higher education programs (or college), g is a control function for the influence of running variable, and ν_i is an exogenous error term.

In the second stage, I estimate the relationship of interest:

$$PolPart_i = \alpha_2 + \beta_2 \cdot Education_i + h(T_i - \tau) + \eta_i \tag{3}$$

where h is the control function and η_i an exogenous error term. The effect of interest is β_2, which is interpreted as the effect of higher education (or college) for the individuals that enrolled because they scored barely above the cutoff, and would have not otherwise, the compliers.

[32] Estimations with the Calonico et al. (2014) optimal bandwidth gives qualitatively the same results.

4 Results Part I. Evidence From Administrative Data

4.1 Exogenous Variation in Education

I start describing the exogenous variation in education. At the loan eligibility cutoff, we observe an important variation in college enrollment and enrollment in any higher education program. The upper graph in Figure 1 shows that pre-selected students, scoring at least 475 points, are twice as likely to enroll in college as those who barely missed the cutoff, i.e., enrollment goes from 18% below the cutoff to 36% above. This is also the case for enrollment in any higher education institution shown in the bottom graph in Figure 1, i.e., students that are barely below the cutoff, thus not eligible for loans for college, imperfectly substitute college with vocational or technical programs, but there is still an increase in any higher education enrollment of about ten percentage points. Table 1 shows regression equivalent results for the variation on education at the cutoff. They show that these results are robust to specification and the inclusion of covariates (with t-stats between 20 and 30).

Students could retake the PSU test once a year, hence they may manipulate the position around the cutoff if they retake the test and succeed scoring above the cutoff. This is problematic, since after a few years there may be no change in enrollment at the cutoff. Panel B in Table 1 shows the estimation for an indicator variable that takes the value of one if a student enroll in any of the three years of available information (from 2007 to 2009), called "ever enrollment". Panel B shows that the jump in enrollment at the cutoff decreases from 18 to 13 percentage points for college, and from 11 to 6 percentage points for any higher education, however the effects are highly significant (in the worst case the t-stat is around 9).

Appendix B shows two tests to confirm the validity of the RD design. First, I show that students are not able to perfectly manipulate their score to benefit from the loans. Figure 5 shows that the density function of PSU scores is continuous at the cutoff (McCrary (2008)).[33] Secondly, Table 11 shows that all baseline characteristics are balance at the cutoff, showing that, arguably, the only element that changes is access to loans for any institution and therefore higher education enrollment. This evidence shows that higher education and college enrollment change as good

[33]Manipulation is also free from the influence of graders or teachers, because the test has only multiple choice question graded by a photo optical device.

as randomly for students at the cutoff, enabling the a reliable estimation of the causal effect of higher education on political participation.

One important aspect of the variation at the cutoff is the fact that higher education is changing in the quantity and quality dimensions. At the cutoff, the number of students that enroll in any higher education program increase significantly, but also, the composition of programs chosen by students move towards more students going to college. The quantity and the quality dimensions should affect participation in the same direction.

4.2 Effects on Registration to Vote

Given the important changes in higher education and college enrollment at the cutoff, the effect on registration to vote can be summarized with the effect of loan eligibility at the threshold, i.e., the reduced form estimation. Figure 2 shows it in its graphical form. The top figure shows registration to vote for the whole PSU support, and in the bottom, a close up for a of 100 PSU points around the cutoff.[34] Both figures show that the relationship is precisely zero.

Table 2 confirms this result. It shows the estimation of equation (1) for different specifications for the control function, f. In columns (1) and (2), f is linear and the estimation is restricted to students that are not more than 44 point away of the cutoff.[35] In columns (3) to (4), f is a fourth order polynomial for the whole PSU support. Even columns add pre-determined covariates. All columns show that crossing the threshold does not imply a change in registration to vote.[36] These results allow to reject that the effect is larger than 0.5% with 95% confidence. This leads to the conclusion that college and higher education in general does not imply a higher political participation in this context.

Appendix C shows the 2SLS regressions of the effects of higher education on registration to vote. College and any higher education are instrumented using the indicator whether student i score at least the loan eligibility cutoff. Again the effects are statistically zero, though not as precise as in the reduced form.

[34] The window of 100 PSU-points corresponds to one standard deviation.

[35] The bandwidth corresponds to Imbens and Kalyanaraman (2012) optimal bandwidth calculation using the edge Kernel. A uniform kernel gives a slightly higher bandwidth, but results do not differ significantly.

[36] The relation is actually slightly negative.

4.3 Heterogeneity

Finding no effect for the overall population could be explained if two groups are being affected in the opposite direction. To rule out this possibility, I check if the effect is different than zero for different sub-populations. I split the sample by the variables that are often cited in the literature as the second most important factor influencing political participation: income.[37,38] Because the majority of students have not yet graduated from higher education in 2009 (the year for electoral data), and therefore do not have their own income source, I will explore heterogeneous effects by family income; specifically, by the classification in income quintiles done by the tax authority. Family income at the time they take the PSU test is a good proxy of the social background students have.

First, I show in Table 3 that the loan cutoff creates strong exogenous variation in higher education for all eligible income quintiles (the poorest four). For the poorest income quintile the effects are the strongest. The probability of college enrollment went from 14.5% (not shown) to 34.5% (a relative effect of 138%). For any higher education program, scoring at least the cutoff implies an increment of 13.5 percentage points in the enrollment rate (a relative increase of 27%, from the baseline enrollment of 49,5%, not shown). These effects are smaller for other eligible income quintiles until the increment is not significant for the (not eligible) richest quintile.

Despite the important changes in higher education, Figure 3 shows that the effects are precisely zero for all sub-populations. Table 4 shows the effects in its regression form for three specifications, confirming that the effects are zero. For the poorest income quintile, it can be rejected a reduced form effect larger 0.7% in the three specifications. For the richest eligible income quintile (the fourth), in the worst case, it can be rejected a reduced form effect of 2.6%. Appendix C show the 2SLS estimations, with the same conclusions, but less precision.

5 Results Part II. Evidence from a Survey

To explore other forms of political participation, political knowledge and to address the issue of overreporting in self-reported information, I use survey data specially collected for this purpose.[39]

[37]The most important being education. See Schlozman (2002).

[38]Appendix D shows the effects for gender, which are essentially the same.

[39]The survey was performed by the Universidad Católica de Concepción, who sent the emails and merged the information to the administrative data from the PSU process, to the enrollment in higher education, and to the

The survey was implemented in October of 2012 to the students that participated in the college admission processes between 2007 and 2009,[40] hence, between four and six years after taking the PSU for the first time.

About ten thousand students responded the survey. Accounting only the emails that were actually opened, the sample implies a response rate of about 40%.[41] There are many reasons that can explain why some students did not open the invitation. First, emails addresses were self-reported by students between 2007 and 2009, and perhaps they were not in use in 2012. Secondly, the invitation email may have gone directly to spam folders because contained words such as "invitation" and "survey" that raise red flags on server security.[42]

To check whether the sample of respondents constitutes a representative sample of the population, I perform two tests: first, I show a test of balance of pre-determined covariates between respondents and not respondents across the cutoff. Second, I show the change in higher education enrollment at the cutoff for the sample of respondents, to compare it with the one for the overall population.

To compare respondents to the overall population, Table 5 shows the estimation of equation (1) fully interacted with an indicator whether the students answered the survey, and where the dependent covariate is each pre-determined variable. Among non-respondents, Column (1) shows the average value for students barely below the cutoff and Column (2) shows the change for those who crossed the threshold. Column (4) shows how respondents differ from the population below the cutoff, and Column (6) how they above. Columns (3), (5) and (7) show respective standard errors.

Two conclusions emerge. First, columns (4) and (5) show that respondents are almost identical to the population. Only household size and type of high school are not balanced, however, the number of characteristics that are not balanced goes in line with type I error.[43] Second,

electoral data. This University is part of the Council of Rectors of the Chilean Universities, who managed the PSU test every year and owns the data.

[40]The full set of questions and the invitation email can be found in Appendix E.

[41]A tracker was added in the email to know whether the email was opened. However, only 13% of the invitations were actually opened, which gives an unconditional response rate of 5%.

[42]Additionally, the invitation contained words such as "raffle", "Ipad", and "tablet", which were used to motivate participation. Finally, the sender email address was not previously known by the students, therefore many students may have erased for security.

[43]Of the 13 tests, is it expected to reject a true hypothesis of no difference 0.65 and 1.3 times with 5% and 10% significance respectively.

columns (6) ad (7) show that, among respondents, almost all characteristics are balanced across the cutoff, and therefore, surveyed students, above and below the threshold, can be used as a good counterfactual to test the causal effect of education on political participation measures.

To further compare surveyed and no surveyed, Table 6 shows the change in higher education enrollment at the cutoff for different specifications. Both measures of higher education enrollment increase strongly at the cutoff, and are numerically and statistically the same as the ones shown in Table 1. As a consequence, the surveyed sample can be considered a good representative of the population of interest.

The survey asked individuals in 2012 to recall if they were registered in order to vote at the end of 2009, which can be potentially problematic if they misrecall. Because the 2009 election was the last one when individuals had to be registered to vote, and the first presidential election where student could have registered (they turned 18 at least between 2007 and 2009) there were no more instances to register between 2009 and the time of the survey to get confused, Moreover, if they registered to vote, they did it with the clear intention to participate in the election. I assume that misrecall is not an important issue in this context.

Table 7 shows the estimations for several measures of political participation, attitudes towards democracy and equality, and self-reported information about political information. The sample is restricted to students scoring 69 points around the loan cutoff, to highlight the characteristics of the group that allows the identification.[44] Columns (1) to (3) shows the correlation between the variables and college, columns (4) to (6), the correlation with respect to any higher education enrollment, and column (7) to (9) show the reduced form estimation (equation (1)). For almost all the variables, the correlation with both measures of higher education is very strong, confirming the findings from the political science literature.[45]

More importantly, the reduced estimation shows that, despite the important change in higher education enrollment that happens at the cutoff, there is no change in (almost) any political participation variable at the cutoff, confirming the results shown previously with administrative data. The only two variables that appear significant at the 10% level are self-reported registration

[44]This bandwidth corresponds to the optimal bandwidth calculation using Imbens and Kalyanaraman (2012) (Edge Kernel) for the sample size of the survey. The effects are qualitative the same using other specifications.

[45]Column (3) shows t-test instead of standard errors to highlight the strength of the relation. Most of the time, the t-stat is above three for college.

and watching politics on TV. Since we have the true registration to vote, the significance in self-reported registration is not a problem and will be reviewed in detail in the following sections. Finding significance at the 10% for "*watching politics on TV*" is in line with type I error.[46]

With this evidence I conclude that higher education is not causing political participation or sophistication.

6 Potential Mechanisms

The survey data offers the opportunity to test two hypotheses that explains the link between education and political participation, which are usually mentioned in the literature as the most important. The cognitive hypothesis argues that (higher) education gives individual the tools to understand the political process. The social network hypothesis, argues that individuals in higher education interact with more politically active agents that either provide information about the political process or social norms that push individuals to participate. I test both of these hypotheses in this section.

6.1 Channel: Education Provides Political Knowledge

If higher education equips students with tools to better understand the political process (the cognitive hypothesis),[47] students should show higher knowledge on elements of the political process four to six years after entering higher education (at the time of the survey). Moreover, a test of political knowledge offers a good opportunity to examine whether education affects the quality of the political decisions (see Milligan et al. (2004) for a discussion). Assuming that more informed individuals make better decisions, the knowledge test shed light on whether education affects political sophistication. The survey questionnaire allows testing how much individuals know about key elements of the political process, giving providing evidence that is free from overreporting bias, an issue usually raised when using survey data.

The political knowledge test in the survey included questions about the position of specific politicians, about the length on the political term for different level of elected representatives and

[46]The reduced form results are robust to the different specifications and the inclusion of covariates. Results upon request.

[47]See for example Wolfinger and Rosenstone (1980) and (Galston (2001).

about names of elected politicians in the district surveyed live. Because, educated individuals may have incentives to look for the correct answers in the internet (there were no monitoring or time restrictions to respond in the survey), the survey also asked whether individuals were able to recognize pictures of important politicians (see the complete survey in Appendix E).

Table 8 shows results for the knowledge test. Panel A and B show the correlation with the two measures of higher education used throughout the paper and Panel C shows reduced form regressions, as in equation (1). Correlations with knowledge are strong for both measures of higher education, especially for college.

However, Panel C shows that the correlation is not causal. Column (1) shows the summary for all 14 questions, indicating that effects larger than 1% can be rejected. The only variable that appears significantly different than zero is the ability of naming their own city mayor (at the 10% level of significance).

The survey was performed one month before the 2012 major election, thus, if voters acquire knowledge a few weeks before the election, it may seems that educated may have an advantage, but that advantage disappear quickly over time. Unfortunately, the analysis of education and its interaction with the time to vote cannot be further explored, and goes beyond the scope of the paper.

6.2 Channel: Overreporting

A few papers argue that the way through education correlates with political participation is through the response to social norms (see for example Silver et al. (1986) and Bernstein et al. (2001)). More educated are more aware of these norms, thus more prone to overreport. The survey allows me to test this claim using a reliable measure of overreporting and a exogenous variation in education.[48]

To motivate the analysis, first, I show the true correlation between registration to vote and higher education and the correlation calculated with survey data. Table 9 shows the (true) correlation is .05 for college and 0.1 for any higher education (columns (1) and (2) respectively). On the other hand, the correlations for the same group[49] using survey self-reported data are

[48] Overreporting is defined as an indicator whether the student reported to be registered while the administrative data indicated the contrary.

[49] Restricting the sample to high school graduates between 18 and 26 years of age.

.15 and .2, three and two times higher than the true, respectively.[50] This evidence shows that overreporting is an issue for survey data analysis, also in this context. Additionally, the table shows that, for the population that allows identification of the causal effects (students around the cutoff), the correlation between higher education and registration is strong.

Previously, Table 7 shows that self-reported registration was weakly caused by higher education, in this section, I argue that this effect is mainly driven by overreporting. Table 10 shows reduced form estimation for three specifications, and 2SLS estimations where college or any higher education enrollment are instrumented by the indicator of scoring at least the cutoff.

The first column, using a lineal control function around the I&K optimal bandwidth (98 points), shows that overreporting increase about 2 percentage points at the cutoff and is significant at the 1% level. The following two columns confirm this finding for different specifications (for a fourth order polynomial over the whole PSU support and linear over the previously used bandwidth, 69 PSU points respectively). Interpreting the results for the set compliers is difficult because there are two elements that change at the cutoff: the number of students enrolling in post-secondary education (any higher education enrollment), and the quality of the programs that are chosen by the students (more students in college). Overreporting may occur because more students feel the pressure to conform to social norms, regardless the type of programs, or, because college student are more affected either because peer pressure (college students are more likely to interact with registered individuals), and/or, at college, some students receive civic education directly. If the effect is driven by the number of individuals continuing into higher education, the complier group corresponds to the individuals that enroll in any higher education program, in which case I instrument enrollment in any higher education program using the indicator for scoring at least the cutoff. The 2SLS estimator indicates that students are 15 percentage points more likely to overreport when they enter any higher education.

On the other hand, if the effects are driven by the quality of education induced by college, the cutoff instruments college enrollment, and the 2SLS estimate suggests that college cause an increment of 10 percentage points in the probability of overreporting.

The effect on overreporting may appear high, but can be explained, in part, because the elec-

[50] I use data from the survey Encuesta Nacional de Opinión Pública from Diego Portales University. Surveys from 2007 to 2011 are pooled together.

tion is a presidential one. Górecki (2011) suggest that presidential elections generate higher social norms increasing the probability of overreporting. Unfortunately, I cannot test if overreporting is lower in non-presidential ones, because, the registration to vote was eliminated after 2010 and the exogenous variation in education did not exist before 2007.[51] However, since the web survey is responded without the influence of the surveyor, it that may reduce the level of overreporting from the presence of surveyors (see for example, Silver et al. (1986) and Holbrook and Krosnick (2010)). As long as these biases are distributed smoothly across the loan cutoff, they do not affect the internal validity of the estimates.

7 Conclusion

The relationship between education and political participation has been broadly explored in the literature, but only few papers have examined the causal channels. To deal with endogeneity, I used a regressions discontinuity design caused by the eligibility criteria of two higher education loans in Chile. The two programs require that students score above a cutoff in the national admission test for college, the PSU test, which is taken by 80% of high school graduates each year. Students that do not meet this requirement can only access to loans in technical institutions.

The eligibility criteria induce a jump in the enrollment rate of college and higher education in general (see Solis (2012)). Detailed individual level data on higher education, for the universe of individuals enrolling in all institutions in the country, was merged with administrative records of voting registration to have a measure of participation that do not suffers from overreporting and response bias.

The estimation indicates that, despite the sharp change in the probability of enrollment in higher education and college, there is no change in registration to vote. Moreover, I find that this results is persistent to many sub-samples, grouped by family income (using tax authority classification), by gender, or other background characteristics.

To explore other forms of political participation, political knowledge, attitudes, and difference in the information sets, I collected survey data for a representative sample of the population and combined it with the same exogenous variation on higher education and the administrative data.

[51]The loan programs with the common requirement started in 2006, but in that year, some mistakes is the assignment rule prevented high changes in higher education enrollment.

I found that higher education does not cause changes in these different forms of political participation (measured as participation on demonstrations, civic organizations, etc.), better attitudes towards democracy, or self-reported measures of information.

Additionally, I tested two often cited mechanism through which education is linked to political participation in the literature, i.e. the cognitive and the social network hypotheses. I found that after four and six year of higher education, students have no more knowledge than their counterfactuals. However, I found higher education does cause overreporting on registration to vote confirming previous findings indicating that educated individuals feel higher pressure to participate and to engage in social desirable behavior. This result shows that using self-reported vote data, potentially distorts the relative effects of education on political participation.

References

Acemoglu, D., S. Johnson, J. A. Robinson, and P. Yared (2005): "From Education to Democracy?" *The American Economic Review*, 95, 44–49.

Barro, R. (1999): "Determinants of Democracy," *Journal of Political Economy*, 107, S158–S183.

Berinsky, A. J. and G. S. Lenz (2010): "Education and Political Participation: Exploring the Causal Link," *Political Behavior*, 33, 357–373.

Bernstein, R., A. Chadha, and R. Montjoy (2001): "Overreporting Voting: Why It Happens and Why It Matters," *Public Opinion Quarterly*, 65, 22–44.

Bobba, M. and D. Coviello (2007): "Weak instruments and weak identification, in estimating the effects of education, on democracy," *Economics Letters*, 96, 301–306.

Borgonovi, F., B. d'Hombres, and B. Hoskins (2010): "Voter Turnout, Information Acquisition and Education: Evidence from 15 European Countries," *The B.E. Journal of Economic Analysis & Policy*, 10.

Calonico, S., M. D. Cattaneo, and R. Titiunik (2014): "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs," *Econometrica*, 82, 2295–2326.

Campante, F. R. and D. Chor (2012): "Why was the Arab World Poised for Revolution? Schooling, Economic Opportunities, and the Arab Spring," *The Journal of Economic Perspectives*, 26, 167–187.

Campbell, A., P. E. Converse, W. E. Miller, and D. E. Stokes (1960): "The American Voter," *Wiley, New York*.

Castelló-Climent, A. (2008): "On the distribution of education and democracy," *Journal of Development Economics*, 87, 179–190.

Chevalier, A. and O. Doyle (2012): "Schooling and Voter Turnout: Is There an American Exception?" SSRN Scholarly Paper ID 2056729, Social Science Research Network, Rochester, NY.

Dee, T. S. (2004): "Are there civic returns to education?" *Journal of Public Economics*, 88, 1697–1720.

Friedman, M. (1962): *Capitalism and Freedom*, Chicago: University of Chicago Press.

Funk, P. (2010): "Social Incentives and Voter Turnout: Evidence from the Swiss Mail Ballot System," *Journal of the European Economic Association*, 8, 1077–1103.

Galston, W. A. (2001): "Political Knowledge, Political Engagement, and Civic Education," *Annual Review of Political Science*, 4, 217–234.

Glaeser, E. L., G. A. M. Ponzetto, and A. Shleifer (2007): "Why does democracy need education?" *Journal of Economic Growth*, 12, 77–99.

Górecki, M. A. (2011): "Electoral Salience and Vote Overreporting: Another Look at the Problem of Validity in Voter Turnout Studies," *International Journal of Public Opinion Research*, 23, 544–557.

Hahn, J., P. Todd, and W. V. d. Klaauw (2001): "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design," *Econometrica*, 69, 201–209.

Hanushek, E. A. (2002): "Chapter 30 Publicly provided education," Elsevier, vol. 4, 2045–2141.

Hillygus, D. S. (2005): "The Missing Link: Exploring the Relationship Between Higher Education and Political Engagement," *Political Behavior*, 27, 25–47.

Holbrook, A. L. and J. A. Krosnick (2010): "Social desirability bias in voter turnout reports Tests using the item count technique," *Public Opinion Quarterly*, 74, 37–67.

Imbens, G. and K. Kalyanaraman (2012): "Optimal Bandwidth Choice for the Regression Discontinuity Estimator," *The Review of Economic Studies*, 79, 933–959.

Imbens, G. W. and J. D. Angrist (1994): "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 62, 467–475.

Imbens, G. W. and T. Lemieux (2008): "Regression discontinuity designs: A guide to practice," *Journal of Econometrics*, 142, 615–635.

Kam, C. D. and C. L. Palmer (2008): "Reconsidering the Effects of Education on Political Participation," *The Journal of Politics*, 70, 612–631.

Karp, J. A. and D. Brockington (2005): "Social Desirability and Response Validity: A Comparative Analysis of Overreporting Voter Turnout in Five Countries," *Journal of Politics*, 67, 825–840.

Lee, D. S. (2008): "Randomized experiments from non-random selection in U.S. House elections," *Journal of Econometrics*, 142, 675–697.

Lee, D. S. and T. Lemieux (2010): "Regression Discontinuity Designs in Economics," *Journal of Economic Literature*, 48, 281–355.

Lipset, S. M. (1959): "Some Social Requisites of Democracy: Economic Development and Political Legitimacy," *The American Political Science Review*, 53, 69–105.

Mazumder, B. (2008): "Does Education Improve Health? A Reexamination of the Evidence from Compulsory Schooling Laws," SSRN Scholarly Paper ID 1134064, Social Science Research Network, Rochester, NY.

McCrary, J. (2008): "Manipulation of the running variable in the regression discontinuity design: A density test," *Journal of Econometrics*, 142, 698–714.

Milligan, K., E. Moretti, and P. Oreopoulos (2004): "Does education improve citizenship? Evidence from the United States and the United Kingdom," *Journal of Public Economics*, 88, 1667–1695.

Moretti, E. (2004): "Chapter 51 Human capital externalities in cities," Elsevier, vol. 4 of *Cities and Geography*, 2243–2291.

Nie, N. H., J. Junn, and K. Stehlik-Barry (1996): *Education and democratic citizenship in America*, University of Chicago Press.

Oreopoulos, P. and K. G. Salvanes (2011): "Priceless: The Nonpecuniary Benefits of Schooling," *The Journal of Economic Perspectives*, 25, 159–184.

Pelkonen, P. (2010): "Length of compulsory education and voter turnout-evidence from a staged reform," *Public Choice*, 150, 51–75.

Rolando, R., J. Salamanca, and I. Rubilar (2010): "Duración Real de las Carreras y Programas. Descripción y Análisis de la Cohorte de Titulados y Graduados 2007," Tech. rep., Higher Education División, Ministry of Education of Chile.

Schlozman, K. (2002): "Citizen Participation in America: What do we know? Why do we care?" in *Political Science: The State of the Discipline*, ed. by I. Katznelson and H. Milner, W. W. Norton \& Company, Inc.

Siedler, T. (2010): "Schooling and Citizenship in a Young Democracy: Evidence from Postwar Germany," *Scandinavian Journal of Economics*, 112, 315–338.

Silver, B. D., B. A. Anderson, and P. R. Abramson (1986): "Who Overreports Voting?" *The American Political Science Review*, 80, 613–624.

Solis, A. (2012): "Credit Access and College Enrollment," Working Paper, Department of Economics, Uppsala University, No. 2013:12.

Sondheimer, R. M. and D. P. Green (2010): "Using Experiments to Estimate the Effects of Education on Voter Turnout," *American Journal of Political Science*, 54, 174–189.

Stephens, M. and D.-Y. Yang (2014): "Compulsory Education and the Benefits of Schooling," *The American Economic Review*, 104, 1777–1792.

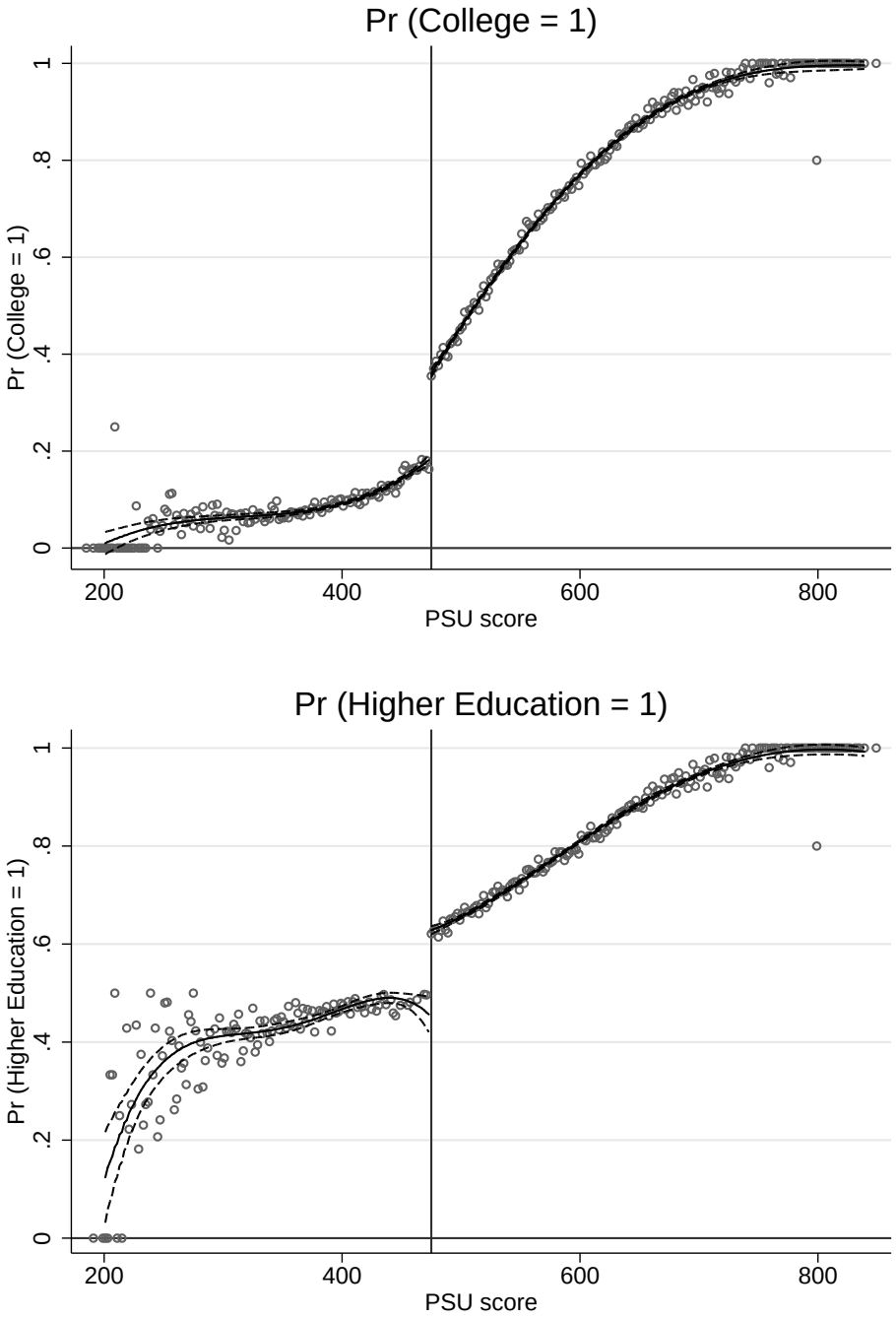
Tenn, S. (2007): "The Effect of Education on Voter Turnout," *Political Analysis*, 15, 446–464.

Wantchekon, L., M. KlaÅ¡nja, and N. Novta (2015): "Education and Human Capital Externalities: Evidence from Colonial Benin*," *The Quarterly Journal of Economics*, qjv004.

Wolfinger, R. E. and S. J. Rosenstone (1980): *Who votes?*, vol. 22, Yale University Press.

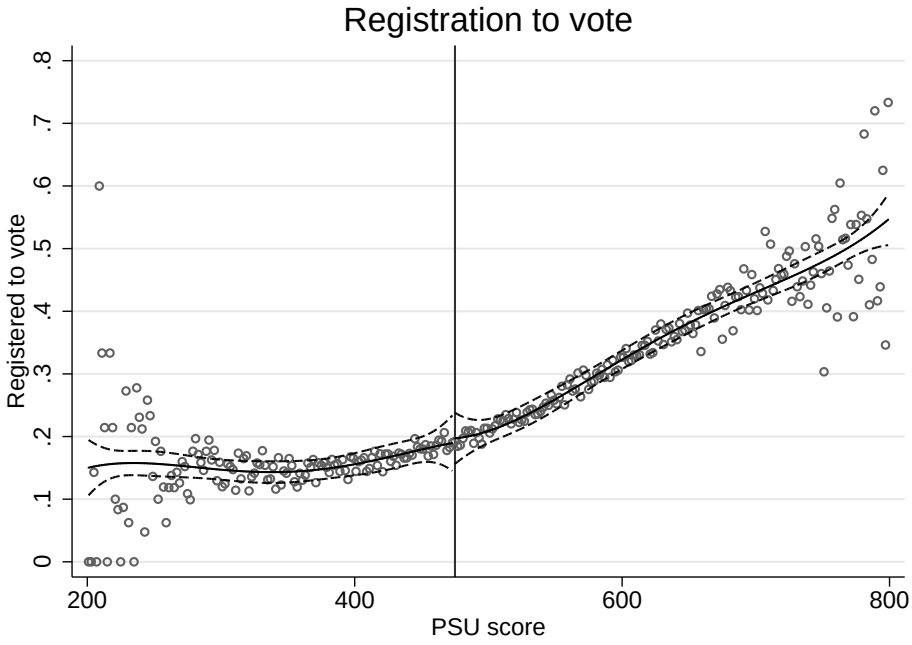
8 Tables & Figures

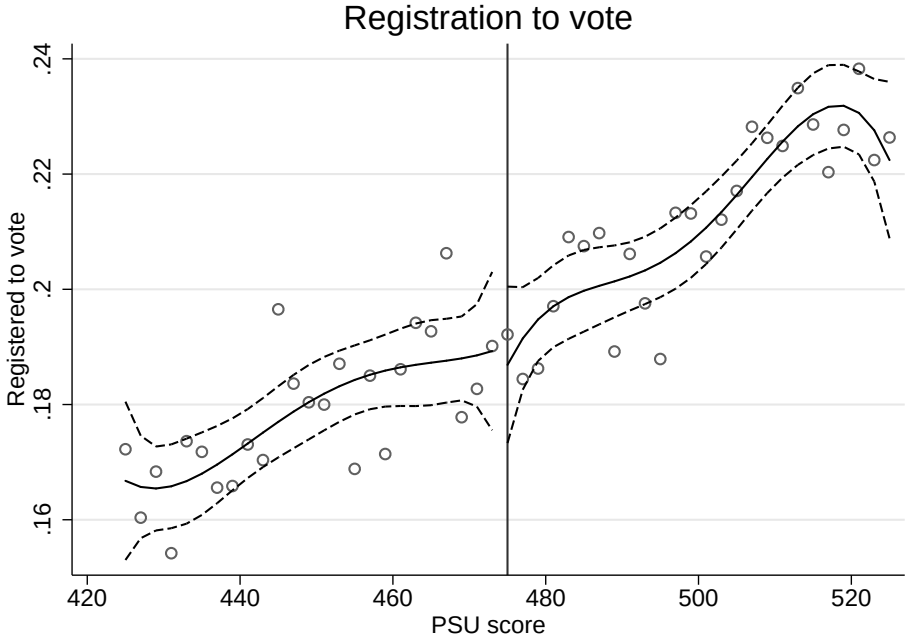
Figure 1: Reduced forms of the higher education measures.



Note: Figures depicts first stages. The figure on the right shows the effect of scoring at least the loan cutoff in college enrollment, while the figure on the left shows the effect in any post-secondary program enrollment (vocational, technical and college). Each dot represents average enrollment within bins of 2 PSU points. The figures consider PSU first time takers that applied to benefits and were classified as eligible-for-loans by the tax authority (pre-selected). The vertical line (at 475) corresponds to the loan cutoff. Dashed lines represent fitted values from the estimation of eq. (1) where $f(\cdot)$ is a 4th order polynomial spline, and 95% confidence intervals for each side.

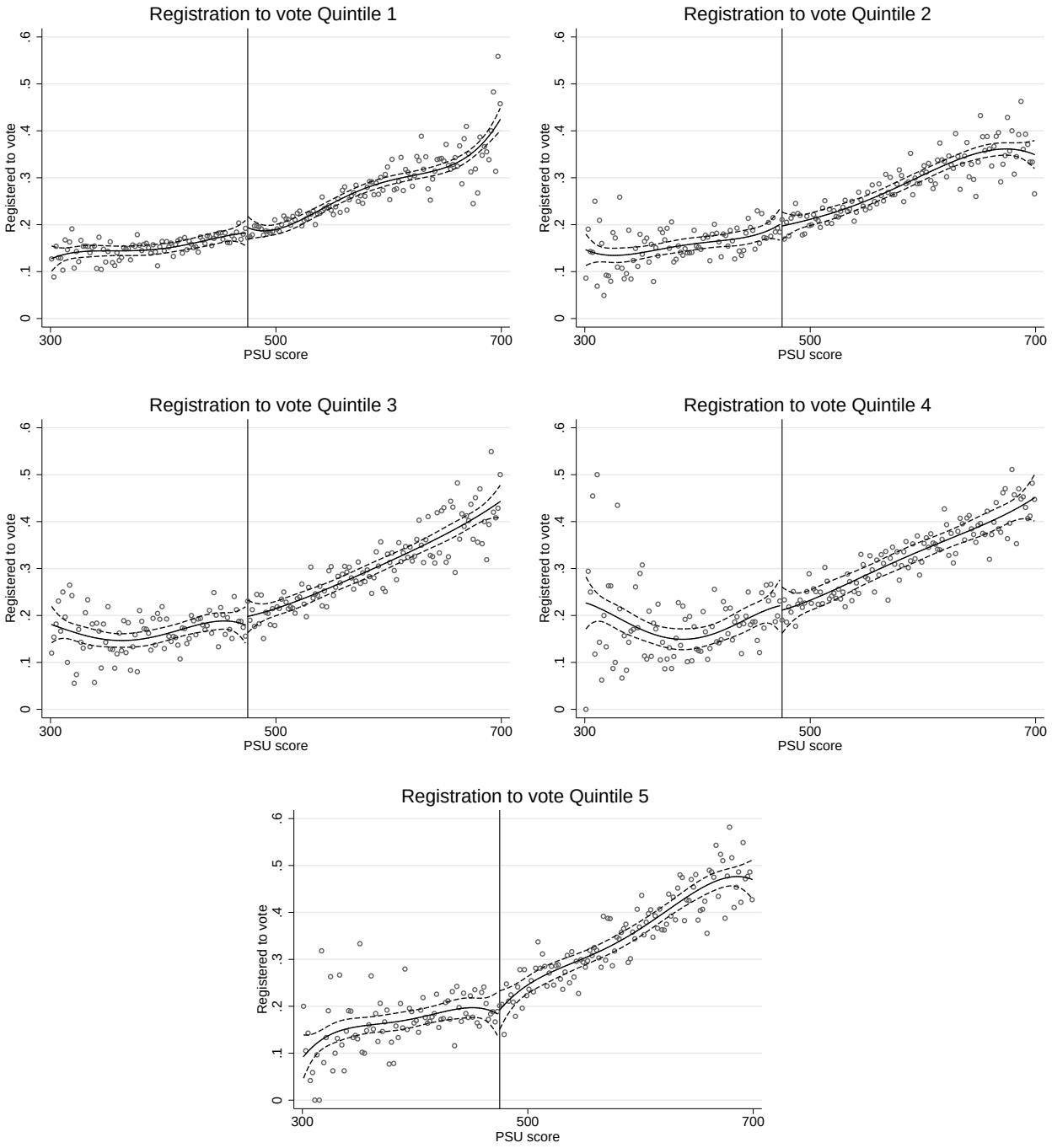
Figure 2: Reduced Form. Registration to vote and in political party on scoring more than the cutoff ($1(PSU_i > 475)$)





Note: Each dot represents average registration to vote within bins of 2 PSU points. The figures consider PSU first time takers that applied to benefits and were classified as eligible-for-loans by the tax authority (pre-selected). The left figure shows the whole range of PSU scores, and the one on the right for a smaller window of 100 points around the eligibility cutoff. The vertical line (at 475) corresponds to the loan cutoff. Dashed lines represent fitted values from the estimation of $PolPart_i = \beta_0 + \beta_1 \cdot \mathbb{1}(T_i \geqslant \tau) + f(T_i) + \nu_i$ where $PolPart_i$ is an indicator whether i is registered to vote, and $\mathbb{1}(T_i \geqslant \tau)$ an indicator whether the student scored at least the cutoff. $f(\cdot)$ is a 4th order polynomial spline controlling for the running variable, and 95% confidence intervals for each side.

Figure 3: Reduced Form by income quintile. Registration to vote.



Note: Each figure shows a different income quintile (income quintile 1 being the poorest). For all figures, each dot represents average registration to vote within bins of 2 PSU points. The figures consider PSU first time takers classified as eligible-for-loans by the tax authority (pre-selected). The vertical line (at 475) corresponds to the loan cutoff. Dashed lines represent fitted values from the estimation of the reduced form equation where $f(\cdot)$ is a 4th order polynomial spline, and 95% confidence intervals for each side.

Table 1: First Stages for College and Any Post-secondary Education.

	𝟙(College)				𝟙(Any Higher Education)			
	Linear		4th Poly		Linear		4th Poly	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Immmediate Enrollment								
$\mathbb{1}(T_i \geqslant \tau)$	.175	.175	.170	.169	.114	.114	.114	.114
	(.006)***	(.006)***	(.006)***	(.006)***	(.007)***	(.007)***	(.007)***	(.007)***
Const.	.224	-.015	.243	.074	.521	.409	.526	.414
	(.007)***	(.014)	(.014)***	(.015)***	(.010)***	(.017)***	(.021)***	(.022)***
Obs.	79348	79254	235801	235552	79348	79254	235801	235552
R^2	.107	.117	.337	.342	.026	.03	.101	.105
Panel B: Ever Enrollment								
$\mathbb{1}(T_i \geqslant \tau)$	.133	.133	.127	.126	.056	.056	.057	.057
	(.007)***	(.007)***	(.007)***	(.007)***	(.006)***	(.006)***	(.007)***	(.007)***
Const.	.337	-.010	.338	.112	.732	.570	.728	.604
	(.005)***	(.015)	(.005)***	(.009)***	(.005)***	(.015)***	(.005)***	(.010)***
Obs.	77288	77288	230653	230653	77288	77288	230653	230653
R^2	.095	.115	.341	.35	.019	.027	.086	.092
Covariates	No	Yes	No	Yes	No	Yes	No	Yes

Note: This table shows estimates of the effect of scoring at least the loan cutoff in higher education enrollment (college, in the first 4 columns, and enrollment in any higher education program in the last 4 columns). Odd (even) columns show estimates without (with) covariates (high school GPA, income quintile reported by the tax authority, self-reported income categories, a female indicator, father and mother education in years, indicator for type of school (voucher, and private - public is the reference), age, work status and married status at the PSU. For each education measure, the first two columns use a linear control function for the running variable, and the next two column a 4th order polynomial. The linear specification restricts the sample to students 44 PSU points around the loan eligibility cutoff (Imbens and Kalyanaraman (2012) optimal bandwidth).

Robust standard error in parenthesis. ***: p-value⩽1%, **: p-value⩽5, *: p-value⩽10%.

Table 2: Reduced Form Estimation on Registration to Vote.

	Lineal		Poly 3rd		Poly 4th	
	Dependent Variable: Registration to Vote					
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{1}(T \geqslant \tau)$	-.005	-.007	-.005	-.007	-.003	-.005
	(.006)	(.006)	(.005)	(.005)	(.006)	(.006)
Const.	.209	-.416	.214	-.402	.211	-.406
	(.008)***	(.028)***	(.009)***	(.017)***	(.016)***	(.022)***
Obs.	79888	79791	235801	235552	235801	235552
R^2	.002	.014	.029	.042	.029	.042
Covariates	No	Yes	No	Yes	No	Yes

Note: Reduced form estimation of the effect of scoring at least the cutoff on registration to vote. Specifically $Registration_i = \beta_0 + \beta_1 \cdot \mathbb{1}(T_i \geqslant \tau) + f(T_i) + X_i \cdot \delta_1 + \nu_i$. To show that the outcome does not relies in the functional form or on the score window used in the estimation, the first two columns uses a linear f (not shown), i.e. $f(T_i) = \gamma_0 \cdot T_i + \gamma_1 \cdot T_i \cdot \mathbb{1}(T_i \geqslant \tau)$ and a window of 44 points around the cutoff. Whereas, f in columns (3) and (4) is a third order, and in columns (5) and (6) a fourth order polynomial. In the last 4 columns the estimation uses the whole sample of students. The outcome is also robust to the inclusion of predetermined covariates. The covariates included are indicators for income quintiles, and self-reported income, age, mother education (in years), father education, indicators for high school of graduation (public, voucher of private), sex, married status, type of health insurance, household size, all variables measured when students took the PSU test for the first time.

Robust standard error in parenthesis. ***: p-value$\leqslant$1%.

Table 3: First stages by income quintile. The effect of scoring at least the cutoff on College Enrollment and Enrollment in any post-secondary programs.

	Lineal	Pol3	Pol4	Lineal	Pol3	Pol4
	(1)	(2)	(3)	(4)	(5)	(6)
	First Stage					
	College			Higher Education		
$\mathbb{1}(T \geqslant \tau) \times q_1$	.200	.192	.198	.135	.136	.135
	(.008)***	(.007)***	(.008)***	(.010)***	(.008)***	(.010)***
$\mathbb{1}(T \geqslant \tau) \times q_2$	.170	.173	.166	.103	.110	.101
	(.013)***	(.011)***	(.014)***	(.015)***	(.013)***	(.016)***
$\mathbb{1}(T \geqslant \tau) \times q_3$	.165	.154	.162	.099	.114	.120
	(.017)***	(.014)***	(.017)***	(.018)***	(.016)***	(.020)***
$\mathbb{1}(T \geqslant \tau) \times q_4$	.070	.072	.066	.051	.064	.045
	(.019)***	(.016)***	(.020)***	(.021)**	(.018)***	(.022)**
$\mathbb{1}(T \geqslant \tau) \times q_5$	.033	.005	.024	.020	-.004	-.002
	(.022)	(.018)	(.022)	(.024)	(.020)	(.025)
Obs.	87050	265606	265606	87050	265606	265606
R^2	.379	.622	.622	.583	.672	.672

Note: First stage estimation for the two measures of higher education: College is an indicator whether i enrolled in any college program (the 3 first columns). Higher education is an indicator whether student i enrolled in a vocational, technical or college program (the last three columns). The estimation correspond to equation (2) fully interacted with indicator variables for each income quintile. Only the interaction between the instrument and the income quintile are shown. The first column in each higher education measure corresponds to a linear specification of the control function, f, the second uses a third order polynomial, and the third column a polynomial of fourth order to show robustness. The linear specification is also restricted to the sample of students scoring at most 44 points away of the cutoff. The polynomial specifications use the whole sample of students.

Robust standard error in parenthesis. ***: p-value$\leqslant$1%, **: p-value$\leqslant$5%.

Table 4: Reduced Form Estimation on Registration to Vote by Income Quintile.

	Lineal		3rd Poly		4th Poly	
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{1}(T \geqslant \tau) \times q_1$	-.008	-.008	-.008	-.008	-.006	-.007
	(.008)	(.008)	(.007)	(.007)	(.008)	(.008)
$\mathbb{1}(T \geqslant \tau) \times q_2$	-.004	-.003	-.00005	.0006	.012	.012
	(.012)	(.012)	(.011)	(.011)	(.013)	(.013)
$\mathbb{1}(T \geqslant \tau) \times q_3$	.020	.019	.015	.014	.005	.004
	(.015)	(.015)	(.014)	(.014)	(.016)	(.016)
$\mathbb{1}(T \geqslant \tau) \times q_4$	-.027	-.028	-.019	-.019	-.020	-.021
	(.017)	(.017)	(.016)	(.016)	(.019)	(.019)
$\mathbb{1}(T \geqslant \tau) \times q_5$	-.005	-.007	.017	.016	.026	.024
	(.019)	(.019)	(.017)	(.017)	(.020)	(.020)
Obs.	84196	84196	259800	259800	259800	259800
R^2	.198	.201	.267	.271	.267	.271
Covariates		X		X		X

Note: Each column show a different specification of the control function f, indicated in the heading. Each regression are equivalent to the estimation of equation (1) fully interacted with dummy for each income quintile. For the linear specification ((1) and (2)), the sample is restricted to the I&K optimal bandwidth (44 PSU-points). For the third and fourth order polynomials the sample includes whole PSU support. Robust standard error in parenthesis.

Table 5: Balance of Covariates for Surveyed students.

	β_0	$\mathbb{1}(T_i \geqslant \tau)$	se	$\mathbb{1}(Survey_i)$	se	$\mathbb{1}(Survey_i)$ $\times\mathbb{1}(T_i \geqslant \tau)$	se
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Income Quintile	1.83	0.017	(0.015)	0.064	(0.08)	-0.029	(0.104)
Self-reported income	1.28	-0.002	(0.007)	0.051	(0.038)	-0.039	(0.049)
1(female)	0.59	0.003	(0.007)	0.048	(0.036)	-0.062	(0.048)
Mother Education (years)	10.66	-0.002	(0.051)	0.175	(0.275)	0.028	(0.345)
Father Education (years)	10.64	-0.008	(0.056)	0.136	(0.325)	-0.071	(0.404)
Household size	4.49	-0.019	(0.026)	0.278	(0.143)*	-0.095	(0.176)
Age at survey	22.77	0.055	(0.022)**	-0.160	(0.106)	-0.041	(0.153)
1(married)	0.01	0.000	(0.002)	-0.012	(0.005)	0.016	(0.008)*
1(work)	0.08	-0.003	(0.004)	-0.017	(0.019)	0.019	(0.025)
1(Public high school)	0.47	0.013	(0.007)*	0.084	(0.038)**	-0.085	(0.049)
1(Voucher high school)	0.51	-0.012	(0.007)	-0.100	(0.038)	0.106	(0.049)**
1(Private high school)	0.01	-0.002	(0.002)	0.012	(0.01)	-0.018	(0.011)
High School GPA	5.60	0.002	(0.006)	-0.010	(0.031)	0.029	(0.041)

Note: Each row show the difference in baseline characteristics among surveyed students, above and below the cutoff, and relative to the whole population. Specifically each row shows the estimation of $y_i = \beta_0 + \beta_1 \cdot \mathbb{1}(T_i \geqslant \tau) + \beta_2 \cdot \mathbb{1}(Survey_i) + \beta_3 \cdot \mathbb{1}(Survey_i) \times \mathbb{1}(T_i \geqslant \tau) + f(T_i) + \vartheta_i$. Where y_i is the covariate in the first column, $\mathbb{1}(T_i \geqslant \tau)$ is an indicator whether student i scored at least the cutoff τ, $\mathbb{1}(Survey_i)$ is an indicator whether student i answered the survey, and f a linear function controlling for the running variable interacted with $\mathbb{1}(T_i \geqslant \tau)$ and $\mathbb{1}(Survey_i)$. The sample is restricted to students 44 around the loan eligibility cutoff. The coefficients on f are not shown.

Robust standard error in parenthesis. ***: p-value$\leqslant$1%, **: p-value$\leqslant$5%, and *: p-value$\leqslant$10%.

Table 6: First Stages on Surveyed Students for College and Higher Education.

	$\mathbb{1}$(College)				$\mathbb{1}$(Vocat., Techn. or College)			
	Lineal		4th Poly		Lineal		4th Poly	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\mathbb{1}(T \geqslant \tau)$	.182	.180	.184	.186	.118	.121	.141	.144
	(.035)***	(.037)***	(.040)***	(.042)***	(.039)***	(.041)***	(.046)***	(.049)***
Const.	.212	-.055	.210	-.038	.549	.501	.524	.402
	(.025)***	(.095)	(.031)***	(.059)	(.031)***	(.093)***	(.040)***	(.063)***
Obs.	2657	2335	8450	7441	2657	2335	8450	7441
R^2	.121	.137	.298	.302	.036	.045	.109	.117
Covariates	No	Yes	No	Yes	No	Yes	No	Yes

Note: Estimates of the relationship between loan cutoff and enrollment in 2 measures of higher education (Enrollment in college, the first 4 columns, and enrollment in vocational, technical or college programs, the last 4 columns) for the students that answered the survey. To show that the change in higher education is not sensitive to specification of f or the inclusion of covariates different columns show different combinations of these two elements. Odd columns show estimates without covariates and even columns include them. The covariates are income quintile reported by the tax authority, self-reported income categories (from 1 to 3), female indicator, father and mother education in years, indicator for type of school (public, voucher, and private), age, work status, married status, and high school GPA. For each education measure, the first two columns use a linear control function for the running variable, and the next two columns a 4th order polynomial. As before, the linear specification restricts the sample to students 44 PSU points around the loan eligibility cutoff (Imbens and Kalyanaraman (2012) optimal bandwidth). Robust standard error in parenthesis. ***: p-value$\leqslant$1%, **: p-value$\leqslant$5%, and *: p-value$\leqslant$10%.

Table 7: OLS, Reduced Form and 2SLS Regressions for Political Participation, and Knowledge and College

		OLS: College			OLS: Higher Education			Reduced Form		
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent Variable	Scale	Level	*College*	t-stat	Level	*H.Educ.*	t-stat	Level	$\mathbb{1}(T_i \geqslant \tau)$	t-stat
A. Political Participation										
Officially Registered to vote	1/0	0.248	0.061	(4.47)***	0.258	0.029	(1.98)**	0.251	0.019	(0.66)
Self-reported registration to vote	1/0	0.258	0.060	(4.32)***	0.270	0.022	(1.52)	0.245	0.047	(1.66)*
Affiliated to a political party	1/0	0.018	0.006	(1.31)	0.018	0.004	(0.86)	0.026	0.004	(0.4)
Participate in demonstrations	1/0	0.454	0.107	(7.07)***	0.481	0.034	(2.07)**	0.458	0.031	(0.96)
Part. demonstration intensity	0-6	1.152	0.406	(7.87)***	1.256	0.129	(2.32)**	1.191	0.051	(0.49)
Participate in civic organization	1/0	0.135	0.040	(3.64)***	0.142	0.018	(1.54)	0.174	-0.032	(1.37)
Part. organization intensity	0-6	0.200	0.068	(3.63)***	0.206	0.038	(1.96)*	0.251	-0.048	(1.25)
B. Attitudes towards Democracy										
Likes politics	1-9 to 0-1	0.280	0.042	(5.04)***	0.287	0.019	(2.11)**	0.262	0.016	(0.94)
Democracy is the best system	1-9 to 0-1	0.605	0.040	(4.92)***	0.608	0.022	(2.41)**	0.612	0.008	(0.46)
Educated are better rulers	1-9 to 0-1	0.438	0.030	(3.52)***	0.440	0.018	(1.96)*	0.443	-0.002	(0.1)
Income should be more equitable	1-9 to 0-1	0.784	0.009	(1.57)	0.787	0.001	(0.14)	0.775	0.017	(1.36)
Taxes should be raised for equity	1-9 to 0-1	0.389	0.026	(2.87)***	0.382	0.029	(2.96)***	0.409	-0.013	(0.7)
More State-own firms	1-9 to 0-1	0.553	0.028	(3.35)***	0.557	0.014	(1.55)	0.552	0.007	(0.38)
Competition is good	1-9 to 0-1	0.640	-0.007	(0.94)	0.624	0.018	(2.23)**	0.646	-0.003	(0.22)
C. Self-reported information										
Knows candidates proposals	1-9 to 0-1	0.409	0.055	(6.07)***	0.413	0.032	(3.25)***	0.409	0.022	(1.15)
Reads politics on newspapers	1-9 to 0-1	0.269	0.043	(5.2)***	0.269	0.029	(3.2)***	0.247	0.025	(1.47)
Watch politics on TV	1-9 to 0-1	0.385	0.042	(4.85)***	0.389	0.024	(2.47)**	0.375	0.034	(1.81)*

Note: Columns (1) to (6) show correlations for two measures of higher education (college and any higher education program) ($Y_i = \beta_0 + \beta_1 \cdot Education_i + \epsilon_i$). Columns (7) to (9) show the reduced form estimation (equation (3)). In Panel B and C (attitudes and self-reported information) answers are in the scale 1one to nine: 1 mean completely disagree and 9 completely agree, and then transformed to the range [0,1]. Scale 1/0 means 1 for yes and 0 for no. "Part. demonstration intensity" corresponds to the sum of six different forms of political demonstrations (street marches, occupying institutions, "cacerolazo" (making noise using a pan), street protests, political meetings, and other demonstrations). The variable "Part. organization intensity" corresponds to the number of organizations where students participate (the intensity of the participation), defined as to the sum of 6 different types or organizations (board of neighbors, sports clubs, student unions, worker unions, religious group, and others organization).

$t - stats$ in parenthesis. ***: p-value$\leqslant$1%, **: p-value$\leqslant$5, *: p-value$\leqslant$10%.

Table 8: Testing the Cognitive Hypothesis

	Overall Political Knowledge	Electoral System test	Knows ministers of government	Recognize politicians faces	Able to name own senator	Able to name own representative	Able to name own mayor
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A. OLS: College Correlation							
College	.007	.041	.014	.038	.074	.070	.031
	(.002)***	(.007)***	(.006)**	(.014)***	(.015)***	(.016)***	(.011)***
Const.	.108	.426	.586	.499	.246	.336	.852
	(.001)***	(.005)***	(.004)***	(.010)***	(.010)***	(.010)***	(.008)***
B. OLS: Higher Education Correlation							
Higher Education	.003	.029	.005	.014	.041	.028	.027
	(.002)**	(.008)***	(.007)	(.015)	(.015)***	(.017)*	(.012)**
Const.	.109	.425	.589	.506	.252	.348	.847
	(.001)***	(.006)***	(.006)***	(.013)***	(.013)***	(.014)***	(.010)***
C. RD: Reduced Form							
$\mathbb{1}(T_i \geqslant \tau)$	.005	.017	.010	.043	.040	.010	.043
	(.003)	(.015)	(.013)	(.029)	(.030)	(.033)	(.025)*
Const.	.107	.427	.575	.497	.250	.343	.825
	(.003)***	(.012)***	(.010)***	(.023)***	(.023)***	(.025)***	(.020)***
Scale	0-14 to 0-1	0-4 to 0-1	0-8 to 0-1	0-2 to 0-1	0/1	0/1	0/1
Obs.	3756	3756	3756	3756	3756	3756	3756

Note: Panel A and B show OLS estimations for two measures of education: college and any higher education, respectively ($Y_i = \beta_0 + \beta_1 \cdot Education_i + \epsilon_i$). Panel C shows the reduced form estimation (equation 1). For the three panels the sample is restricted to students that are 69 PSU points around the loan cutoff (I&K optimal bandwidth) .

Robust standard errors in parenthesis. ***: p-value<1%.

Table 9: Correlation between registration to vote and education. OLS estimation

	Official records			Self-reported in Survey		
	Higher Educ.	College Educ.	Years of Educ.	Higher Educ.	College Educ.	Years of Educ.
	(1)	(2)	(3)	(4)	(5)	(6)
Const.	.199	.184	.181	.139	.149	-.270
	(.001)***	(.001)***	(.001)***	(.014)***	(.012)***	(.062)***
Education	.046	.102	.018	.150	.210	.038
	(.002)***	(.002)***	(.0004)***	(.024)***	(.030)***	(.005)***
Obs.	235801	235801	235801	1130	1130	1128
R^2	.003	.014	.009	.034	.054	.047
Source	Registry	Registry	Registry	UDP	UDP	UDP

Note: Each column shows the OLS estimation of the relationship between registration to vote and education, $Regis_i = \beta_0 + \beta_1 \cdot Education_i + \epsilon_i$, where $Education_i$ is an indicator whether individual i have enrolled in any program from post-secondary education labeled "higher education" (columns (1) and (4)), that have enrolled in college programs ((2) and (5)) or number of years of education ((3) and (6)). The first three columns show the relationship using administrative data, while the last three columns use one of the most important political surveys in the country: Encuesta Nacional de Opinión Pública from the Diego Portales University. Surveys from 2007 to 2011 have been pooled together, and the sample is restricted to individuals between 18 and 26 years old. Higher education is an indicator whether a student has enrolled in any vocational, technical or colprogram (when using registry data) and an indicator for the categories complete/incomplete technical education or college and also graduate school. College is an indicator for students that have enrolled in college programs in the period 2007-2011 (when using registry data), and an indicator for individual that report educational level as complete or incomplete college or graduate school.

Robust standard errors in parenthesis. ***: p-value<1%.

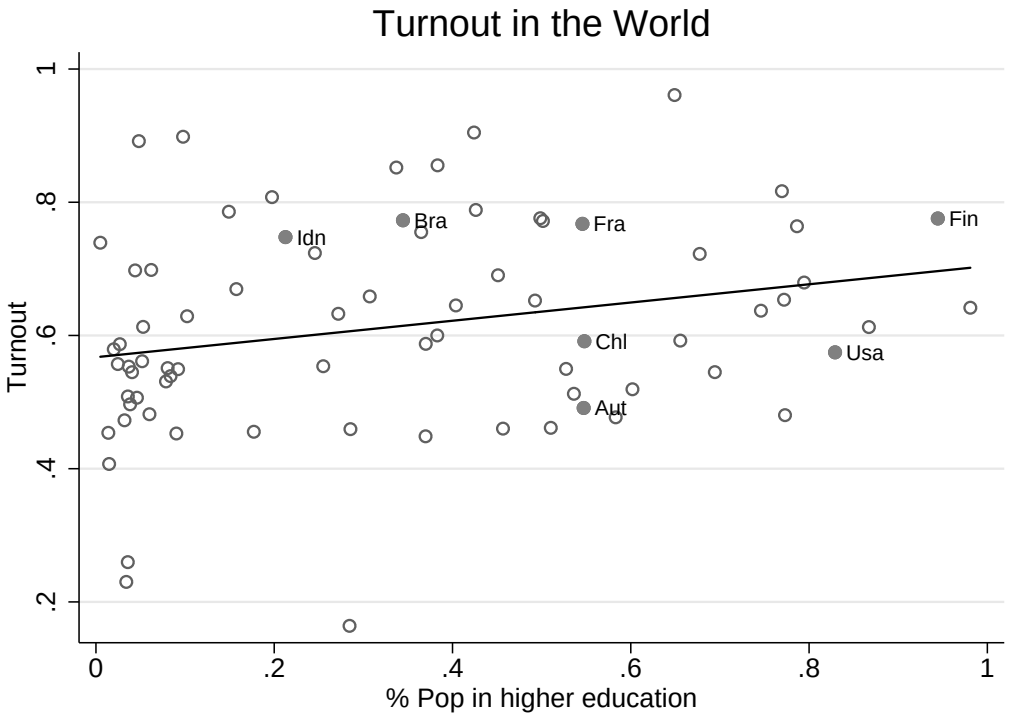
Table 10: Overreporting. Testing the Social Network Hypothesis

	Reduced Form			2SLS	
	(1)	(2)	(3)	(4)	(5)
$\mathbb{1}(T_i \geqslant \tau)$	.022	.027	.019		
	(.008)***	(.011)**	(.008)**		
$\mathbb{1}(college)$				.096	
				(.036)***	
$\mathbb{1}(H.Educ.)$					.165
					(.070)**
Const.	.006	.010	-.013	-.014	-.085
	(.006)	(.008)	(.036)	(.013)	(.044)*
Bandwidth	I&K	All	69	I&K	I&K
Obs.	5166	8422	4493	5166	5166

Note: The first three columns show reduced form estimations for the measure of overreporting (an indicator equal to one if a surveyed answer to be register when the administrative data indicates the contrary). The first and third columns, use a linear control function for the running variable using the I&K optimal bandwidth in the first (98) and 69 points in the third (the bandwidth previously used for the survey results). The second column uses a fourth order polynomial over the whole PSU support. Columns (4) and (5) show 2SLS estimation for college and any higher education enrollment respectively using as instrument the indicator whether the individual scored at least the cutoff in the PSU test.

A Appendix: Chile's Turnout relative to the World

Figure 4: Turnout around the world



Note: Y-axis shows turnout for 52 countries relatively to the share of the population enrolled in higher education. Turnout data from the International Institute for Democracy and Electoral Assistance (IDEA) and population on tertiary education from UNESCO Institute of Statistic Data Centre. It shows that Chile is a country with similar level of political participation (voter turnout specifically) than the US, which has been study extensively in the literature. Also the level of higher education in the population is similar to developed countries such as Austria and France.

B Validity

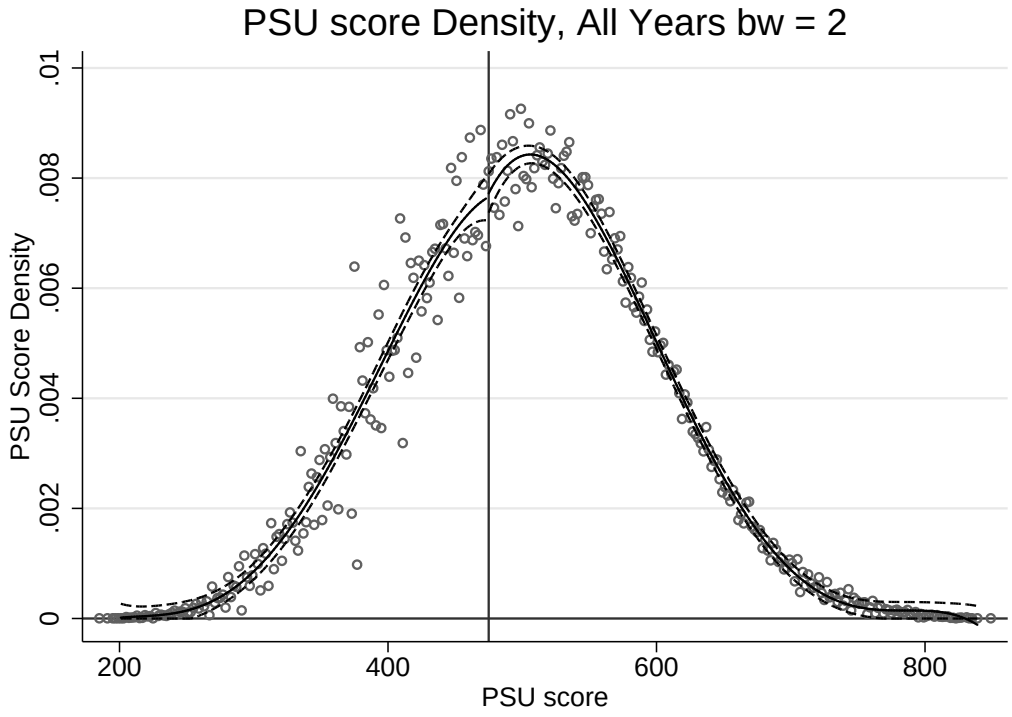
In this Appendix I test if the natural experiment produces exogenous variation in education. Following Imbens and Lemieux (2008), I present two tests. First, I show the score density to show that students cannot perfectly manipulate the PSU score. Second, I present a test for balance of pre-determined covariates to show that in any other dimension students barely above and below the cutoff are comparable.

Figure 5 shows the PSU score density. Dots in the figure correspond to the density for students scoring in bins of 2 points. Moreover, the figure shows fitted values for each side of the cutoff, of a regression using a fourth order polynomial, plus 95% confidence intervals. The figure shows that the density function is statistically continuous at the cutoff, confirming that PSU scores are not subject to manipulation around the cutoff.[52]

Table 11 show estimation of equation (1) for different pre-determined covariates. The first two column uses a liner control function over the I&K optimal badwidth, while columns (3) and (4) use a fourth order polynomial over the whole PSU sport. Columns (1) and (3) show the level for each characteristics, while columns (2) and (4) shows the change for those who score barely above the threshold. For the linear control function, only age appear significantly different at the 5% level of significance. From the fourth order polynomial control function, all variables appear balanced. In expectation, 1.3 tests should reject the true hypothesis o no effects, therefore, I conclude that individuals barely above and below the loan cutoff are comparable and therefore the natural experiment provides as good as random variation to test the causal effect of education on political participation.

[52] Additionally, the test contains only multiple choice question and is graded by an photo optical device, without the intervention of evaluators.

Figure 5: RD for PSU scores frequency distribution.



Note: Each dot represents the density of PSU scores in an interval of 2 points. The sample considers only students who satisfy all requirements to be eligible for college loans and take the PSU immediately after graduating from high school. The vertical line indicates the loan cutoff (475). The three cohorts (2007, 2008 and 2009) of first time takers pooled together. Dashed lines represent fitted values from the estimation of eq. (1) where $f(\cdot)$ is a 4th order polynomial spline, and 95% confidence intervals for each side.

Table 11: Balance of covariates for three groups of students: 1) All students, 2) Only students enrolled in any post-secondary education, and 3) Only college students.

		All		Higher Education		College	
	Level	Linear	4th poly	Linear	4th poly	Linear	4th poly
Dependent Var.:	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Income quintile	2.096	0.016	0.025	-0.027	-0.013	-0.221	-0.210
		(0.015)	(0.016)	(0.02)	(0.022)	(0.034)***	(0.038)***
Self reported Income	1.409	-0.003	0.001	-0.022	-0.017	-0.114	-0.105
		(0.007)	(0.007)	(0.009)**	(0.01)*	(0.016)***	(0.018)***
$\mathbb{1}$(Female)	0.602	0.001	0.006	0.011	0.020	0.016	0.023
		(0.007)	(0.008)	(0.009)	(0.01)*	(0.015)	(0.017)
Mother Education	11.370	-0.005	0.041	-0.008	0.032	-0.455	-0.427
		(0.051)	(0.055)	(0.068)	(0.074)	(0.104)***	(0.118)***
Father Education	11.486	-0.011	0.034	-0.110	-0.077	-0.657	-0.526
		(0.056)	(0.06)	(0.075)	(0.081)	(0.113)***	(0.128)***
Household size	4.458	-0.019	-0.003	-0.010	0.001	0.075	0.088
		(0.026)	(0.028)	(0.035)	(0.038)	(0.054)	(0.061)
Age	17.926	0.043	0.021	0.031	0.014	-0.017	-0.006
		(0.019)**	(0.024)	(0.027)	(0.032)	(0.036)	(0.042)
$\mathbb{1}$(married)	0.005	0.000	-0.001	0.002	0.002	0.006	0.006
		(0.002)	(0.002)	(0.002)	(0.002)	(0.003)**	(0.003)*
$\mathbb{1}$(work)	0.067	-0.003	-0.004	0.000	-0.003	0.004	0.003
		(0.004)	(0.004)	(0.005)	(0.006)	(0.008)	(0.009)
$\mathbb{1}$(Public high school)	0.412	0.011	0.009	0.021	0.009	0.079	0.064
		(0.007)	(0.008)	(0.01)**	(0.01)	(0.015)***	(0.017)***
$\mathbb{1}$(Voucher high school)	0.558	-0.010	-0.010	-0.021	-0.011	-0.066	-0.055
		(0.007)	(0.008)	(0.01)**	(0.01)	(0.015)***	(0.017)***
$\mathbb{1}$(Private high school)	0.024	-0.002	0.000	-0.002	0.000	-0.013	-0.010
		(0.002)	(0.002)	(0.002)	(0.003)	(0.004)***	(0.005)*
High School GPA	5.579	0.003	0.010	0.010	0.014	0.071	0.067
		(0.006)	(0.006)	(0.008)	(0.009)	(0.013)***	(0.015)***
Observations		77,574	230,758	44,944	147,938	23,699	100,624

Note: Estimates for differences on baseline characteristics among students above and below the loan cutoff. Specifically each row corresponds to the estimation of $y_i = \beta_0 + \beta_1 \cdot \mathbb{1}(T_i \geqslant \tau) + f(T_i) + \epsilon_i$. Where y_i is the covariate in the first column, $\mathbb{1}(T_i \geqslant \tau)$ is an indicator whether student i's score (T_i) is at least equal to the loan cutoff (τ), and $f(T_i)$ is a function controlling for the running variable (flexibly at each size of the cutoff) not shown. f is linear in columns (2), (4) and (6) and a 4th order polynomial in columns (3), (5) and (7). Moreover, the linear specification is restricted to students 44 points around the cutoff (the bandwidth based on Imbens and Kalyanaraman (2012)optimal bandwidth calculation). Column (1) shows the level of each covariate for the linear specification. For the following specifications the level is not shown, but they are quite similar to those presented in column (1). Columns (2) and (3) show the estimation using the population of students that took the PSU test. Columns (4) and (5) show estimations when the sample is restricted to students that enrolled in vocational, technical or college programs in the first year. Columns (6) and (7) restrict the sample to students enrolled in college programs. These four last columns show the sorting process of students matriculating in higher education.

Robust standard error in parenthesis. ***: p-value<1%, **: p-value<5, *: p-value<10%.

C 2SLS results

Table 12: The Effect of Education on Voting Registration. 2SLS estimation.

	Lineal		Poly 3rd		Poly 4th	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: College						
$\mathbb{1}(College)$	-.031	-.040	-.032	-.041	-.020	-.030
	(.032)	(.032)	(.029)	(.029)	(.035)	(.035)
Const.	.216	-.420	.222	-.401	.216	-.405
	(.014)***	(.028)***	(.015)***	(.018)***	(.023)***	(.022)***
Obs.	79888	79791	235801	235552	235801	235552
R^2	-.002	.008	.026	.039	.028	.04
Panel B: Any Higher Education						
$\mathbb{1}(Higher\ Educ.)$	-.048	-.061	-.046	-.058	-.029	-.045
	(.049)	(.048)	(.042)	(.041)	(.053)	(.053)
Const.	.234	-.390	.238	-.376	.227	-.385
	(.031)***	(.038)***	(.029)***	(.030)***	(.041)***	(.040)***
Obs.	79888	79791	235801	235552	235801	235552
R^2	-.0009	.009	.027	.038	.028	.04
Covariates		X		X		X

Note: Two stage least square estimation of the effect of higher education and college on registration to vote, specifically equations (2) and (3). Panel A uses as measure of education an the whether student i enrolled in a college program. Panel B, uses an indicator whether i has enrolled in any type of post-secondary program, i.e., vocational, technical or college. Both measures of higher education are instrumented using the loan eligibility cutoff. In the first two columns f (not shown) is linear, and the sample is restricted to a window of 44 points around the cutoff. In columns (3) to (6) f is a polynomial (third order in (3) and (4), and forth in (5) and (6)) using the whole PSU support. Even columns show regressions that include predetermined covariates. The covariates included are: an indicator for income quintile, self-reported income, age, mother education (in years), father education, indicators for high school of graduation (public, voucher of private), gender, married status, type of health insurance, household size. All predetermined variables were measured when students took the PSU test for the first time

Robust standard error in parenthesis. ***: p-value$\leqslant$1%.

C.1 Heterogenity by Family Income

Table 13 shows 2SLS estimation for the relationship between two measures of higher education and political participation by income quintile using three different specifications. Both measures of education uses the same instrument, the indicator function where an individual scored at least the cutoff. The table shows that there is no effect on registration to vote in neither income quintile.

Each cell corresponds to the estimation for one income quintile and one specification sepa-

rately. Only the coefficient of the education measure, its standard deviation and the number of observation in each regression are shown. In most of the cases the relationship is negative, but never significantly different than zero.

Table 13: 2SLS estimation of the effect of education on registration to vote by income quintile.

	Lineal (1)	Polyn. 3rd order (2)	Polyn. 4th order (3)	Lineal (4)	Polyn. 3rd order (5)	Polyn. 4th order (6)
	2SLS Estimation					
	$Education = \mathbb{1}(College_i)$			$Education =$ $\mathbb{1}$(Any Higher Education)		
$Education_i \times q_1$	-0.042	-0.044	-0.032	-0.062	-0.062	-0.046
se	(0.038)	(0.036)	(0.042)	(0.056)	(0.051)	(0.061)
Obs	[41,540]	[110,896]	[110,896]	[41,540]	[110,896]	[110,896]
$Education_i \times q_2$	-0.015	-0.006	0.059	-0.024	-0.009	0.097
se	(0.071)	(0.063)	(0.079)	(0.117)	(0.099)	(0.131)
Obs	[17,379]	[50,873]	[50,873]	[17,379]	[50,873]	[50,873]
$Education_i \times q_3$	0.096	0.079	0.016	0.16	0.106	0.022
se	(0.09)	(0.088)	(0.101)	(0.154)	(0.12)	(0.136)
Obs	[11,584]	[37,794]	[37,794]	[11,584]	[37,794]	[37,794]
$Education_i \times q_4$	-0.362	-0.302	-0.314	-0.495	-0.34	-0.462
se	(0.266)	(0.227)	(0.297)	(0.383)	(0.258)	(0.465)
Obs	[9,385]	[36,238]	[36,238]	[9,385]	[36,238]	[36,238]
$Education_i \times q_5$	-0.032	3.805	1.198	-0.053	-5.22	-16.414
se	(0.584)	(14.517)	(1.369)	(0.95)	(30.54)	(235.212)
Obs	[7,162]	[29,805]	[29,805]	[7,162]	[29,805]	[29,805]

Note: 2SLS estimation for the relationship between register to vote and higher education ($Education_i$). College and any post-secondary program (Vocational, technical or college) are instrumented with an indicator whether student i scores more than the loan eligibility cutoff. For college as measure of higher education in the 3 first columns, and any post-secondary program in the last three columns. Specifically, the estimation corresponds to equations (2) and (3) for each income quintile estimated separately. Only the parameter for the variable $Education_i$ is shown, plus its standard deviation, and the number observations in each regression (in brackets). Columns (1) and (4) use a linear control function, f. Columns (2) and (5) use a third order polynomials and (3) and (4) a fourth order polynomial respectively to show robustness. The linear specification is restricted to the sample of students scoring at most 44 points away of the cutoff. The polynomial specifications use the whole PSU support. Robust standard error in parenthesis.

D Heterogeneity by gender

One important feature explaining political participation is gender (see Schlozman (2002)). Panel A in Table 14 shows the effects on higher education enrollment of crossing the cutoff by gender. As before, Column (1) uses a linear control function, and columns (2) and (3) a third and a fourth order polynomial respectively. The variation in higher education enrollment is a very strong for both measures: college and any higher education program. Women scoring barely below the loan cutoff have a college enrollment rate of 20%, and crossing the threshold implies an increase of 15 percentage points (a 75% relative increase). For men, baseline college enrollment rate is 25%, and for (barely) eligible for loans, increases by 17.1 percentage points. For any higher education enrollment, we observe similar patterns. Crossing the cutoff implies and increase of nine and eleven percentage points for males and females respectively.

Panel B shows reduced form estimation to show that there is no effect on registration to vote. Finally, Panel C shows the fuzzy RD, i.e., 2SLS estimation of the two higher education measures on registration. Both measures of education are instrumented using an indicator for scoring at least the cutoff. This panel confirms that the effect is statistically zero.

Table 14: First stages, reduced forms and 2SLS estimation for registration to vote by gender.

	Lineal	Polyn. 3rd order	Polyn. 4th order	Lineal	Polyn. 3rd order	Polyn. 4th order
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: First Stage					
	$\mathbb{1}(College)$			$\mathbb{1}(Any\ Higher\ Education)$		
$\mathbb{1}(T_i \geqslant \tau) \times \mathbb{1}(male)$	.151	.147	.145	.097	.093	.092
	(.009)***	(.006)***	(.009)***	(.010)***	(.009)***	(.011)***
$\mathbb{1}(T_i \geqslant \tau) \times \mathbb{1}(female)$	.171	.166	.166	.112	.119	.114
	(.008)***	(.007)***	(.008)***	(.009)***	(.008)***	(.009)***
Obs.	87,050	265,606	265,606	87,050	265,606	265,606
R^2	.374	.617	.617	.582	.669	.669
	Panel B: Reduced Form					
	$\mathbb{1}(male)$			$\mathbb{1}(female)$		
$\mathbb{1}(T_i \geqslant \tau) \times \mathbb{1}(\cdot)$	-.009	-.007	-.010	-.003	-.0001	.006
	(.008)	(.007)	(.009)	(.007)	(.006)	(.008)
Obs.	35,779	115,955	115,955	51,271	149,651	149,651
R^2	.345	.640	.641	.370	.596	.597
	Panel C: 2SLS estimation					
	$Education = \mathbb{1}(College_i)$			$Education = \mathbb{1}(Any\ Higher\ Education)$		
$Education_i \times \mathbb{1}(male)$	-0.069	-0.074	-0.097	-0.107	-0.111	-0.149
se	(0.054)	(0.052)	(0.061)	(0.084)	(0.078)	(0.095)
Obs	[32,815]	[102,288]	[102,288]	[32,815]	[102,288]	[102,288]
$Education_i \times \mathbb{1}(female)$	-0.009	-0.004	0.028	-0.014	-0.006	0.041
se	(0.039)	(0.036)	(0.044)	(0.06)	(0.05)	(0.065)
Obs	[47,073]	[133,513]	[133,513]	[47,073]	[133,513]	[133,513]

Note: Panel A shows the first stage by gender, the first three columns for college and the following three for higher education. Panel B shows the reduced form by gender, the three first columns for males and the next three for females. Panel C shows the 2SLS estimation by gender. In this case, for clarity purposes each regression is run separately to show the number of observations involved. Each column corresponds to a different specification for the control function f. In columns (1) and (4) the control function is linear in a window of 44 PSU points. In columns (2) and (5) f is a third order polynomial, and in columns (3) and (6) a fourth order polynomial using the whole range of PSU scores.

Robust standard error in parenthesis. ***: p-value$\leqslant$1%.

E Survey

E.1 Invitation Email

- **Subject:** Survey Invitation
- **Email Body**:

  Dear <***first name***>,

  We are a group of researchers from the Catholic University of Concepción, conducting a study on the Chilean political participation and wanted to ask your help answering a brief questionnaire that takes about 15 minutes.

  In appreciation for your help, you will participate in the raffle of 2 iPads, 1 tablet HP and 5 gift cards of 10 thousand pesos.

  If you have questions do not hesitate to write to info@encuestapsu.cl or visit the website www.encuestapsu.cl

  To answer the questionnaire, click on the following link

  <***Personalized link***> (If the link does not load, copy and paste it into the address bar)

  Thank you and good luck

  The researchers

  *PS: You are contacted to this email because you registered it in the PSU process of the year 200X. If you do not want to receive this type of emails, simply reply an email with the word "**unsubscribe**" in the subject. Sorry for the inconvenience.*

E.2 Political Participation Survey

Page I. Participation in elections

- **1A.** Will you vote in the next elections? Yes / No
- **1B. Were you registered to vote in the presidential election of December 2009? Yes / No**

  If Yes in 1B:

  1C. Did you vote in the first round? Yes / No / do not remember

  1D. Did you vote in the second round? Yes / No / do not remember

1E. Did you vote to elect senators? Yes / No / do not remember

1F. Did you vote to elect representatives? Yes / No / do not remember

1G. Did you vote to elect mayor in your town? Yes / No / do not remember

Page II. Participation in elections

- **2.**

If Yes in 1C: **2A.** Who did you vote for in the first round?

Jorge Arrate / Marco Enríquez-Ominami / Sebastián Piñera / Eduardo Frei / Null / White / I don't want to answer

If Yes in 1D: **2B.** Who did you vote for in the second round?

Sebastián Piñera / Eduardo Frei / Null / White / I don't want to answer

If Yes in 1E, 1F, 1G: **2C.** Do you remember who you voted for ...?

... In the election for senator Yes / No

... In the election for deputy Yes / No

... In the election for mayor Yes / No

2D. About the place you live in. Please Name ...

... a Senator from your district _____ / not remember

... a congressman in your district_____/ not remember

... the mayor of the municipality where you live _____/ not remember

Page III. Participation

- **3.** Are you enrolled in any political party? Yes / No

If Yes in 3.

3B. Which one?

Party for Democracy / National Renewal / Christian Democracy / Communist Party / Independent Democratic Union / Socialist Party / Humanist Party / Other

4. Have you participated in any political demonstration (march, Building occupation, protest, etc.)? Yes / No

If Yes in 4.

4B. What kind of political demonstration has you participated? (Check all that apply)

a. Marchs / b. Occupying colleges or universities / c. Cacerolazo / d. Street Protest / e. Political meetings / f. Other

5. Do you participate in any civic organization? Yes / No

If Yes in 5.

5B. What kind of organization? (Check all that apply)

a. Board of neighbors / b. Sports club / c. Student Unions / d. (Workers) Union / e. Religious Group / f. Other

Page IV. Elections

- **6. Test**

6A. How often is there elections for ...?

A) ... President?

3 years / 4 years / 5 years / 6 years / 7 years / 8 years / Do not know

B) ... senators in your region?

3 years / 4 years / 5 years / 6 years / 7 years / 8 years / Do not know

C) ... deputies in your district?

3 years / 4 years / 5 years / 6 years / 7 years / 8 years / Do not know

6B. How many senators are elected in your senatorial district?

1 / 2 / 3 / 4 / 5 / Do not knowledge

Page V. Do you Know (2/2)...

- **6. Test**

6D. What position is held by the person in the photograph?

a) President of the Chamber of Deputies
b) Foreign Minister
c) Chairman of the Senate
d) Minister of Finance
e) None of the above

6D. What position is held by the person in the photograph?

a) President of the Chamber of Deputies
b) Foreign Minister
c) Chairman of the Senate
d) Comptroller of the Republic
e) None of the above

Page VI. Do you Know (2/2)...

- **7. Test**

  7A. Of the following people. Who are currently ministers in the government? (Check all that are correct)

  Carlos Larrain Peña / Alfredo Moreno Charme / Manuela De la Barra Hoffman / Cristián Larroulet Vignau / Andres Velasco Brañes / Harald Beyer Burgos / Felipe Bulnes Serrano / Joaquín Lavín Infante

  7B. On a scale of 1 to 9, please indicate how much do you agree with the following statements (1 = NEVER - 9 =ALWAYS)

  - I like to talk about politics with my friends: 1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9

  - Vote or not, I'm always informed on candidates' proposals: 1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9

  - I follow politics reading the newspaper: 1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9

  - I follow politics watching the news on TV: 1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9

Page VII. Elections

- **8.** On a scale of 1 to 9, indicate how much do you agree with the following statements: (1 = Strongly Disagree - 9 =Strongly Agree)

  - Democracy is the best system of government : 1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9

  - In general, Better educated people are better rulers: 1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9

  - People´s income should be more equitable: 1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9

  - Taxes should be raised to improve the country's income distribution: 1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9

  - The number of State-own firms should be increased: 1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9

  - Competition is good, stimulates effort and new ideas: 1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9

  - In general, women are better leaders than men: 1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9

[Here questions no related with political participation survey were omitted]

Page VIII. Study and work

- **17.** Have you enrolled in a career in higher education? Yes / No

  If No in 17:

17A. Why Not?

a. I don't care studying / b. I didn't reach the PSU score for the career I wanted / c. I'll prepared to give back the PSU / d. I don't have the financial means / e. I needed to work / f. Another reason

If Yes in 17:

17A. What Program did you enrolled? (If you have enrolled several times, name the most recent)

Type of institution: _____ / Institution Name: _____ / Program Name: _____ / If your career is not on the list, please click here

17B. Did you graduated from that program? Yes / No

If No in 17B: **17C.** Why didn't you graduate?

a. I'm still studying / b. I changed program / c. I failed / d. I Freeze studies / e. I didn't like it / f. I had to leave due to financial problems / g. Another reason

If Yes in 17B: **17C.** What year did you graduated on that program? 2008 2009 2010 2011 2012

Page XII. Study and work

- **18.** Did you work while studying? Yes / No

  If Yes in 18:

  18A. How much did you earn monthly (approximate)?

  a. 0 to 50 000 pesos / b. Between 50 000 and 100 000 / c. Between 100 000 and 200 000 / d. Between 200 000 and 300 000 / e. More than 300 000

  18B. How many months did you earned the previous amount? _____

  18C. What motivates (motivated) you to work while studying?

  a. Meeting the needs of my family / b. To help with the Household expenses / c. To pay fees and tuition / d. To pay other expenses associated with the studies (the board, transportation, photocopying, books etc.) / e. For my entertainment / f. To know what is to work / g. Another reason /

Page XIII. Labor Situation Today (Last Page)

- **19.** Are you working now? Yes / No

  19A. WHAT IS YOUR MONTHLY INCOME TODAY? Monthly Income $: _____ (Enter only numbers and dots. Example: 600.000)

  If you prefer not to report your income, we offer the following:

  a. 0, (I'm still a student) / b. Less than 182 000 (minimum wage) / c. Between 182 000 and 300 000 / d. Between 300 000 and 500 000 / e. Between 500 000 and 700 000 / f. Between

700,000 and 1 million / g. Between 1 million and 1 million 300 thousand / h. More than 1 million 300 thousand

[end of the survey]