Student loans: Liquidity constraint and higher education in South Africa

Marc Gurgand, Adrien Lorenceau, Thomas Mélonio

WORKING PAPER N° 2011 - 20

Student loans: Liquidity constraint and higher education in South Africa

Marc Gurgand

Adrien Lorenceau

Thomas Mélonio

JEL Codes: I23, I24, I25, H52, G21
Keywords: Education, university, credit constraint, regression discontinuity

Student loans: Liquidity constraint and higher education in South Africa*

Marc Gurgand (Paris School of Economics, Crest and J-Pal)
Adrien Lorenceau (Paris School of Economics)
Thomas Mélonio (Agence française de développement)

May 4, 2011

Abstract

Empirical evidence that access to higher education is constrained by credit availability is limited and usually indirect. This paper provides direct evidence by comparing university enrollment rates of South African potential students, depending on whether they get a loan or not to cover their registration fees, in a context where such fees are high. We use matched individual data from both a credit institution (Eduloan) and the Department of Education. Based on a regression discontinuity design using the fact that loans are granted according to a credit score threshold, we can estimate the causal impact of loan obtainment. We find that the credit constraint is substantial, as it decreases the enrollment rate into higher education by more than 20 percentage points in a population of student loan applicants.

*We are very grateful to Herman Steyn and all Eduloan staff for providing invaluable data and many precisions and advices regarding the South African higher education system and their work on the field; and to Jean Skene and Dr. Qhobela from the Department of Education (DoE) for providing us with the HEMIS data. Thanks to Jocelyn Vass, Fabian Arends, Mariette Visser from HSRC for very valuable discussions and advice, and to Geneviève Michaud for assistance with the data. Special thanks also to the AFD team in South Africa. We are grateful to Clément de Chaisemartin, Eric Maurin and seminar participants at Pompeu Fabra, Paris School of Economics, European Development Network Workshop, EALE/SOLE (2010) and Econometric Society World Congress (2010) for useful comments. The authors gratefully acknowledge financial help of the Agence Française de Développement (AFD). The views expressed herein are those of the authors only.

1 Introduction

Whereas primary education is almost universal in South Africa, and secondary schooling has a very large outreach, higher education has become a severe issue in this emerging country. Enrollment is about 15%, a low figure at this level of development. Limited access is strongly concentrated on the black and colored population and, generally, on the poor. Meanwhile, wage returns to university degrees are high. This involves both efficiency and equity considerations that stand high on the political agenda.

Credit constraint seems a natural interpretation of such a situation. Higher education is costly, both in terms of direct and opportunity cost, and modest people may be unable to borrow against future income if credit markets are imperfect. Although such imperfection is likely, its extent remains debatable in a rather highly financialized country. Moreover, the observed stylized facts can also be explained by other types of deprivation, such that the poor happen to lack academic qualification, or taste, for university studies.

If credit constraint is a major issue, then a relevant policy is to encourage the provision of student loans. This paper assesses the impact of a private company supported by international donors, Eduloan, that provides short term loans to pay for university fees. In South Africa, average university fee is equivalent to the average monthly wage and, in many institutions, it can be 2 to 5 times that much. Our estimation is based on a population willing to enroll into a university and asking Eduloan for a loan. We compare actual enrollment of individuals that obtain such loan with individuals that don't. Identification of a causal effect is based on the observation that a credit score threshold is used by Eduloan to decide on loan grants: following the regression discontinuity approach, we can compare otherwise similar individuals with and without a loan, on either side of the discontinuity.

We could match application and client data from Eduloan with individual data on university students from the South African Ministry of Education (*Hemis* data). Therefore, we observe both loan requests, loan grants and subsequent enrollment and graduation for a large sample of individuals. With this data, we can show that loan access substantially increases the probability to actually enroll by 20 to 25 points, representing a 50% increase. The effect tends to be even stronger for poorer families.

This result can be interpreted as a positive impact evaluation of Eduloan. But it also brings new and straightforward evidence that liquidity constraint is a significant obstacle to higher education. Given the high level of fees, even a short term loan can affect enrollment, as many people (in particular among the poorest potential students) obviously have no alternative when they are denied a loan from Eduloan. In this unique setup, we can both show that the rest of the market exposes individuals to liquidity constraint and that this constraint has a large (and quantified) impact on enrollment decisions.

Beyond the Eduloan case, this paper contributes to the literature on borrowing constraint and the demand for education. Proving the existence of such constraint and measuring its extent has proved a difficult task, and the literature has followed indirect routes. To emphasize the source of the difficulty, think of demand for education S in the standard Beckerian framework as a function of potential wages and interest rate: $S(w(.), r)$. Credit market imperfection imply that individuals face interest rates higher than the market rate, and decreasing with assets or parental income; or that they face a limit to their debt that is also a function of their current resources. Demand for education would then have the form $S(w(.), r(I), d(I))$ where I is a measure of family income and d is maximum debt. Contrasting these two demand functions is hard because $r(I)$ is rarely observed, the market interest rate r is empirically difficult to determine, and debt, if observed, could well be an optimal, not a constrained, amount of debt.

A first strand in the literature has estimated the causal effect of parental income I on education level S. Some authors, for instance Acemoglu and Pishke (2001) or Maurin (2002), claim that there is a positive effect. But Cameron and Heckman (2001) and Carneiro and Heckman (2002) argue that such link only reflects the impact of cultural traits or very early investment during childhood. At any rate, the reduced form demand function with credit constraint is undistinguishable from $S(w(.), r, I)$, a demand function with perfect credit markets but a consumption motive. This approach thus exhibits a credit constraint only if education is believed to be a pure investment good.

Another approach is based on the *discount-rate bias*, thus labeled after Lang and Ruud's (1986) and Lang's (1993) estimation of idiosyncratic discount rates. Card (2001) basically takes one's marginal return to schooling to be an estimate of one's value of $r(I)$. He argues that, for some instruments for schooling in a wage equation, marginal rates of returns are estimated over a population potentially constrained by liquidity. Because, with such instruments, estimated returns are much higher than OLS returns, this could be evidence that $r(I)$ is indeed higher for individuals of modest origin. Cameron and Taber (2004) develop the argument further using a model where only the credit market for human capital is imperfect. In this model, the credit constraint only applies to the direct cost of education. They reconsider Card's argument in this context and estimate a structural model of the form $S(w(.), r, C \times r(I), I)$, where C is the amount of direct cost: this interaction allows to differentiate the effect of $r(I)$ from I. They find no evidence of a credit constraint.

Attanasio and Kaufmann (2009) have recently taken a different route. They claim that subjectively expected wages at various schooling levels ($\widetilde{w}(.)$) are the relevant argument in the demand function $S(\widetilde{w}(.), r)$. As such, $\widetilde{w}(.)$ can be observed, simply by asking people. Schooling should increase with expected returns. If schooling demand is constrained by some binding level of debt $d(I)$, however, then this relationship no longer holds. They do find that expected returns are correlated with actual schooling for the richer part of their sample, but not for the poorer, which seems to indicate that the poor are credit-constrained.

A few other papers, such as Keane and Wolpin (2001), Brown et al. (2009) and Lochner and Monge-Naranjo (2011), rely on structural or calibrated models. Generally, there is little agreement over the existence and importance of credit constraints. The literature is inconclusive and strongly focussed on the developed world. Moreover, the set of empirical methods implemented are extremely *indirect*, in order to circumvent a basic observability issue. In contrast, this paper takes a very *direct* approach, using a quasi-experiment over loan provision. This is probably the most straightforward way to document this issue. If the loan reduces (at least part of) a credit constraint, then it should increase higher education enrollment. On the opposite, if credit constraint is not binding, the loan may well increase individual welfare, but not enrollment.

To our knowledge, the only two papers to take a similar route are Canton and Blom (2004) and Stinebrickner and Stinebrickner (2008). The latter ask American students the hypothetical question whether they would like to take a loan at a fair interest rate. Although this is close to our test as a thought experiment, constrained students are identified based on a subjective question, which may be very different from actually providing or not providing a loan. Furthermore, they do not estimate the effects on attendance. The former use data on actual loan provision in Mexico. However, they cannot measure enrollment impact because all of their population is already enrolled. They estimate impacts on academic performance instead, but they are exposed to strong selectivity bias if loan provision is also a determinant of enrollment.

In this paper, we have an actual quasi-experiment over loan grants and we observe a sample of potential students, some of whom will eventually not enter university. This is a unique setup to provide evidence on credit constraints. Because the loans that we observe are short term, we must make the distinction between "liquidity constraints", for those who would have a sufficient income to enroll, in present value, but may lack the savings to pay for the tuition fees upfront, and "solvency constraints", for those who would need to increase their income after their studies to reimburse their loan. Although the latter would arguably affect a larger share of the population, we show that even a pure liquidity constraint has important consequences for enrollment.

In the following section, we present the Eduloan scheme in the general South African context and show that, although other access to loans is available to some segments of the population (the poorest and the richest), most of the individuals willing to enroll at university may have limited access to credit. We insist that the high level of fees in this country makes even a loan limited to fee payment a potentially important option. We also sketch a model of Eduloan client behavior, in order to clarify the interpretation of the estimated parameter. Section 3 describes the data. Section 4 presents the results and section 5 shows several robustness check using different sources of identification and samples. Finally Section 6 discusses the results, notably their interpretation and external validity. We then conclude.

2 The Eduloan scheme in the South African context

Since the end of the apartheid regime, in 1994, the higher education system has experienced profound changes. The government faced a challenging trade-off: improve access to the historically disadvantaged people while ensuring the development of the educational system along with international standards. The latter has led to the reorganization of the public institutions in three types: the Universities, the Universities of Technology and the Comprehensive Universities (providing both general and vocational qualifications). Distance learning represents more than one third of total enrollment.

However, whereas primary education is universal and secondary enrollment is more than 90%, enrollment in higher education is only around 15%, among which 60% are Black Africans although they represent 80% of the population. Moreover, the graduation rate is extremely low, between 15% and 20% depending on the qualification level and population group (Department of Education, 2009). In this context, access to higher education, especially at the advantage of the historically disadvantaged, remains an issue that is high on the South

African political agenda.

In contrast, wage returns to higher education seem to be very high: Branson et al. (2009) and Keswell and Poswell (2004) argue that marginal returns to schooling are increasing with schooling level and are as high as 50% per year at the tertiary level. Altogether, this set of facts - low attendance and high return - is compatible with some forms of constraint in access to higher education.

An obvious source of constraint could be the "shared cost" principle adopted by the South African higher education system: since private returns to tertiary education are high, "users" are asked to finance it partially. As a result, tuition fees represent about 25% of the higher education budget. In 2004 (the begining of our sample period) they amounted to ZAR 5,251 millions (Stumpf et al., 2008), for 744,000 students. The yearly average fee is thus about ZAR 7,000,[1] with in fact substantial variation between institutions: it is not unusual that required fees are between ZAR 15,000 and ZAR 35,000, especially in contact education (as opposed to distance education).[2] Those fees are to be compared to the average monthly wage which is arount ZAR 7,500 in this period (Statistics South Africa, 2006) or to the annual GDP per capita at ZAR 36,000 in 2006.[3] In the presence of liquidity constraint, such fees could well explain low enrollment and low graduation in spite of high returns.

In order to empower the historically disadvantaged people and increase participation to higher education for the poorest, the government has implemented a contingent loan program (NSFAS). The loans are granted on the basis of a means test. They are to be reimbursed only when the student is employed and the instalments depends on her salary; moreover, 40% of the loan can be converted into a bursary depending on the student's academic results. In 2004, the amounts lent ranged between ZAR 2,000 and ZAR 25,000, and the program benefitted 15% of the students in public institutions (Stumpf et al., 2008), out

[1]Accounting for inflation, this is about 1,200 current US dollars.

[2]Social Surveys, 2009.

[3]Relative to GDP per capita, the ZAR 7,000 fee would be comparable to a US$ 9,500 average fee in the US.

of which 98% are historically disadvantaged.

In the South African financial context, the NSFAS is the main opportunity for poor students to finance their education. Commercial banks constitute an alternative source of financing as they offer student loans as well (Social Surveys, 2009). However, the requirements for a loan approval are such that likely only the wealthiest families will use this option. Informal money lenders also exists, but they charge very high interest rates (40 to 50% on an annual basis). In the light of this financial environment Eduloan holds a very specific market position.

2.1 Eduloan

Eduloan is a private financial company created in the mid 1990's that receives support from international donors, essentially in the form of guarantees for the loans taken by Eduloan from national banks. It provides loans to cover registration fees for individuals planning to enroll at a (public or private) university in South Africa. The position of Eduloan in the student's loan market seems to be between the NSFAS and the commercial banks. It targets middle to upper-middle income households most of whom would not be eligible to the NSFAS but may not be wealthy enough to get funded by commercial banks. This is thus a population likely to face borrowing constraints.

Eduloan provides short to medium term loans (typically 12 to 24 months) at a moderate rate (around 1% above the prime rate, which is the reference rate for households). In order to be eligible borrowers must be employed and have a minimum level of income. In addition, the instalment must not exceed 25% of the monthly salary. Clients can borrow for education undertaken by themselves or by their relatives.

Whether the loan is granted or not also depends on a credit score, called the *Empirica score*. It is computed by a credit bureau based on a national-wide banking history. Although the algorithm is not made public, we know it does include information such as amounts owed, the number of credit cards, delinquency and numerous other variables. The final decision to grant a loan to an applicant is in great part dependent on the applicant's Empirica score being above a threshold, which is not public and that we cannot reveal for confidentiality reasons (the threshold will be normalized to zero). The Empirica will thus be our forcing variable for the regression discontinuity identification strategy. Individuals are unaware of their score and it is very unlikely that they can manipulate its value in the neighborhood of the threshold (that they don't even know and we don't report here).

Loan application works as follows. Eduloan has an office on most public university campuses. A student must first decide on the university and the qualification she wishes to attend. Once the university has accepted her application and provided her with the corresponding quotation fee, she can directly apply to Eduloan to cover part or all of the fees. If her loan request is accepted, Eduloan pays registration fees directly to the university. Notice also that the student may ask for additional loans during the year if needed. The important feature for us is that choice of a university is a prerequisite for loan demand and loans are necessarily provided for that university, because of direct payment. This will allow us to restrict most of our analysis to students who requested a loan in order to enter a public university: they cannot receive a loan that they could use to pay for a different university or for consumption.

2.2 The parameter of interest

In such a context, it is important to clearly describe the parameter that we can estimate. The fact that loans are short term and must be repaid during school, and that potential students have an explicit education project before loan access is revealed, are important and specific features. The following simple model clarifies the extent to which the impact of this specific program can be taken to reveal liquidity constraints.

Let us describe the intertemporal utility of an agent who borrows at Eduloan. We assume the agent has access to resources I every period while studying. Existence of such income is a requirement. It can be the income of her parents, relatives or spouse or her own income if she keeps working while studying. If she decides to enroll at University, she will have to pay a fee f. As mentioned above, for many persons, f may not be negligible with respect to I. Consider 3 periods. In period 1, the agent decides to register or not. If she does, she has I for her living and has to pay f. In order to pay the fees f she can borrow d from Eduloan or use I in any proportion. In the second period, she has a new income I and she must repay the loan, if any. In the third period, she receives her new wage as a more educated person w^H. In order to make very clear that the loan is short term, we assume that repayment must occur before the agent actually receives her wage w^H. The agent has to solve the following program:

$$\max \qquad u(c_1) + \delta u(c_2) + \delta^2 u(c_3)$$

$$\begin{cases} c_1 &= I + d - f \\ c_2 &= I - rd \\ c_3 &= w^H \\ d &\leq \overline{d} \\ d &\geq 0 \end{cases}$$

where δ is a subjective discount factor, r the interest rate on the loan, $\overline{d}$ the maximum amount that can be borrowed from Eduloan, and c_i the consumption in period $i = 1, 2, 3$. In general, the reduced form utility from this program will be some function $V^H(w^H, I, r, f, \overline{d})$ with $\overline{d} = f$ because Eduloan offers a loan that can cover no more than schooling fees f. If the agent had decided not to enroll in higher education, she would earn w^L every period and her intertemporal utility would be of the form $V^L(w^L)$. Eduloan clients, which we observe, thus have several characteristics: they are willing to enroll if they do obtain a loan from Eduloan: $V^H(w^H, I, r, f, \overline{d} = f) > V^L(w^L)$; also, their utilities and incomes are such that they are willing to take a loan $d > 0$. Whatever parameter we estimate is valid only for that specific population.

Evaluating the impact of Eduloan as a scheme amounts to comparing enrollment outcomes when access to the scheme is available and when it is not. If, in the absence of Eduloan, the same person would have access to a commercial bank instead and could borrow a maximum amount $\overline{d}'$, then her intertemporal utility if enrolled would be $V^H(w^H, I, r, f, \overline{d}')$.[4] The proportion of people who asked for a loan from Eduloan and, everything else equal, enroll when they are granted the loan and do not enroll otherwise is thus:

$$P[V^H(w^H, I, r, f, \overline{d}') < V^L(w^L) | V^H(w^H, I, r, f, \overline{d} = f) > V^L(w^L)]$$

and this is the parameter we can estimate if we observe higher education enrollment of similar people that, for arguably exogenous reasons, are or are not granted the Eduloan loan they asked for.

If positive, this parameter contains two pieces of information. The first is that, for a set of individuals, $\overline{d}' < f$: it implies that, but for Eduloan, there is a borrowing constraint in the South African financial market such that those people cannot borrow at least the full amount of fees.[5] In our setup, liquidity constraint is obviously evaluated with respect to Eduloan. Eduloan is by no means financial market perfection: it offers low interest rate, but short term and limited amounts. But this estimation can reveal that, for some individuals, borrowing capacity is even more limited.

The second information is that the constraint $\overline{d}' < f$ is binding for that proportion of people. There could be liquidity constraints, but people enroll in higher education however, because they'd rather cut very strongly their current consumption for instance. When the objective is to increase enrollment in higher education rather than welfare of the students, unbinding liquidity constraint is of limited importance. We only identify the extent of *binding* liquidity constraint, but this is most important on policy grounds.

This paper uses discontinuity in the Empirica score that determines access to Eduloan loan in order to compare similar people with and without a loan.

[4]The argument also runs if we define liquidity constraint as having access to a higher interest rate $r' > r$.
[5]It is easy to show that V^H is non-decreasing with $\overline{d}$.

The loan impact on enrollment is thus an evaluation of the usefulness of the scheme, in the tradition of public policy evaluation. If the market was highly competitive and many banks were willing to lend to the same people under the same conditions, Eduloan would bring no value added at the margin, and donors supporting the scheme would need to know that. On the contrary, if there is an impact, it implies that Eduloan decreases the level of liquidity constraint, and liquidity constraint is indeed a reason why some people may not enter university. From the scheme evaluation, we learn something more general for which, as already mentioned, we have no such direct evidence in the literature.

We can quantify the consequences of such constraints on a specific population however: people willing to enter higher education provided they get a loan. To this extent, we probably underestimate the impact of liquidity constraint: some persons would need a long term loan or a loan that covers more than registration fees in order to engage in higher education. We do not have that population in the data. On another hand, we do not expect credit constraint to be a relevant issue for people who do not think of entering university, because of taste or capacity. Sorting out this latter population from individuals who are liquidity constrained is an issue in the literature. We directly exclude this population here.

3 Data

The data used in this paper have two sources. The first is customer data from Eduloan, that describe loan application and obtention. This is necessary in order to compute a "treatment" variable over a population of interest. The second is provided by the Ministry of Education and identifies the students entering a public higher education institution, thus the outcome variable. These two data sets are merged using national identification numbers.

3.1 Eduloan data

As a private credit company, Eduloan maintains customer files on both the whole set of applicants and on their actual customers. They have provided us with two data sets. The first one contains information on Eduloan applicants between 2004 and 2008. The key variables are the Empirica score, the national identification number of the student (who is not necessarily the applicant when parents borrow for their children) and the application date. In addition, the files include characteristics on the applicant such as the borrower's net salary, the institution she applies for, the loan amount requested, her age and so on. The second data set contains actual customers, i.e. the applicants whose loan application was accepted and received a loan. Again, the key variables for our purpose are the national identification number and the agreement date.

In the first dataset, we can observe several application dates per applicant and per year. It can either be duplicate administrative records for the same request or individuals who actually apply for more than one loan over the year. When a loan has been granted, we have no direct information over which application it corresponds to. Because our outcome (university enrollment) is a yearly event, it is enough for us to know if, for some year, some applications were sent and some loans were obtained.

In most of the empirical analysis, we use data from 2004 to 2007, because this is the period over which the threshold Empirica value set by Eduloan's internal procedures does generate a discontinuity on loan grants. During this period, the threshold remained unchanged. In 2008, Eduloan's activities were strongly impacted by the credit crunch that followed the financial crisis, and the threshold had much less explanatory power. We use the 2008 data only for a robustness analysis.

3.2 HEMIS data

The second source of data originates from the Ministry of Education, which manages its public subsidies to higher education institutions based on enrollment figures. The Higher Education Management Information System (HEMIS) has therefore been created to collect accurate individual data on each and every student entering the public higher education system. The data contains information on all the courses and qualifications undertaken by a student throughout her studies in the public institutions. This includes the name of the institution, the type of courses or qualifications, educational credits completed among those taken, whether the student is in contact or distance mode, etc.

As this data contains the student national identification number, it can be matched with the Eduloan applicant and client data. Our database is unique, starting with a list of more than 15,000 applications for a loan at Eduloan, complemented with systematic information on whether they obtained a loan from Eduloan and whether they enrolled and completed their credits in a public higher education institution during the relevant year.

3.3 Data constraints

The major limitation of this data stems from the fact that Hemis files only contain information on students entering *public* higher education institution. Therefore, we do not know whether individuals who applied to private higher education institutions eventually enrolled. In South Africa, the private higher education sector is quite developed with around seventy noticeable institutions.

Fortunately, loans are granted in order to pay fees to a specific institution and they are paid by Eduloan directly to that institution. When a loan has been required for a public institution (what we will call thereafter "Hemis perimeter"), then we know if granting the loan has indeed increased the chances that the applicant actually enrolled. Our data contains a variable for the type of institution the student asked a loan for. However, this variable is not properly filled for about 18% of observations. When information is available, a large majority of students (80%) applied to a public institution as compared to 20% to a private one.

Our baseline analysis will be restricted to applicants within the Hemis perimeter, excluding loan requests for a private or unknown institution. We will check that such sample selection is independent from having an Empirica score on either side of the threshold. Because this is verified, the sample restriction has no implication for internal consistency; but it affects external validity. In our robustness analysis, we will include the sample with unknown institution and show that we can then estimate a lower bound to the effect on Hemis perimeter individuals. But we will never have any claim over the population that wishes to enter private institutions.

The other technical difficulty is to match dates between applications, loans granted and enrollment. The academic year is the civil year in South Africa. The norm is that students register in January and ask for a loan right away: 55% of our application dates are in January or February, and 62% if we include March. But some office treatment may take time and some students may ask for help to pay additional fees or a second fees instalment later on, so that some additional applications appear all over the year. We keep only one observation per student and per year. We consider that loans requested year t have been granted whenever the same student has put one or more applications during year t and has received a loan during that year t.

There is an ambiguity, however, when loan requests are posted late in the year and a loan is granted at the beginning of the next year. We are uncertain whether it is meant to pay for late fees or if the loan demand is in provision for the coming year. We are thus unsure whether this very demand has been accepted and we don't know if we should relate this demand to enrollment the current year or the next. As a result, our baseline estimation excludes observations for which the only application of the year was posted in November or December (we then keep 86% of our sample). We will test, as a robustness check, that results are not sensitive to inclusion of those observations. Of course, it is still possible to misallocate an application to an academic year, despite dropping the late applications.

Finally, it is worth mentioning that we had to drop some observations for which the national identification number was missing or obviously incorrect. Also, some individuals with no credit history did not have an Empirica score: they are excluded from the whole analysis.

3.4 Descriptive Statistics

Table 1 presents our sample for the years 2004 to 2007, on which most of the analysis is based. Each observation corresponds to a loan request for a student and for a given year. As explained above, when the earliest demand was placed in November or December, the loan/student/year observation is not included in the baseline sample.

The table shows the characteristics of the student, of the loan request and of enrollment in a public university, if any. The figures are presented separately for individuals who requested a loan to pay for the fees of a public university (Hemis perimeter), of a private institution, and for whom this information was missing in the data. We also split the sample between requests that were granted a loan and those that were turned down.

It is important to note that the average student age is high, typically around 27. This is mostly explained by the fact that a large share of the students are the borrowers themselves who, by Eduloan rules, have to be working with a regular income and a payslip. A substantial share of the sample population are employees willing to upgrade their qualification in order to get access to better paid jobs, and not just parents borrowing for their children's education. This is common practice in South Africa, where the largest university in the country (UNISA) is dedicated to distance education.

Borrowers report wages that are relatively high by South African standards:

the average monthly wage is between ZAR 6,000 and 7,500. This is to be compared with the average wage in the population in formal employment, which is around ZAR 7,500 in this period (Statistics South Africa, 2006). Given that wages are usually log-normally distributed and accounting also for the existence of informal employment, it is very likely that our population of borrowers are somewhat above the median wage. Therefore, our sample can be regarded as a collection of potential students from middle-class South African households, although probably not the most well-off. This is precisely the population that we expect to pursue higher education (having passed high school and been accepted academically in a university), but may face liquidity constraint doing so. As a matter of fact, requested loan values represent on average one to two monthly wages, an amount that households may find difficult to make available up-front, but are likely to repay over 12 to 24 months. This is also a reminder that our sample is obviously not representative of South African population, but may correspond to those for which liquidity is a binding constraint.

Overall, Eduloan gratifies 42% of requests. Loans are granted more often to borrowers that report higher wages (by about ZAR 1,000 in all samples). However, the proportion of males, the proportion of students who are themselves the borrower and their age are not quite different depending on loan status.

When we consider loans requested for a public university, 75% of students who were granted a loan ended up actually enrolled according to the Hemis database, whereas only 53% did so when they did not obtain a loan from Eduloan. As a result, a naive estimation of loan impact would be an additional 22 points, or 41% increase in enrollment rate. The fact that a quarter of students who planned to enroll and did get a loan are not eventually enrolled has no single explanation. One obvious possibility is that they changed their minds, faced unexpected constraints, did not obtain complementary resources, etc. A very likely explaination is that they dropped out early in the year: indeed Hemis data do not include early drop outs, and we already mentioned that drop out rates are huge in South Africa. If students drop out in spite of the loan, this will indeed reduce the estimated loan impact. Finally, we cannot exclude mistyped ID numbers or other sorts of mismatches, such that some enrolled persons are treated as non-enrolled or the reverse. However, given that enrollment is an explained variable and we will use an instrument that must be independent from such measurement errors in the outcome, this should only come at the cost of statistical precision.

Among students actually enrolled in a public university, loan status is only associated with a small difference in the number of courses they register at, and in the number of credits they obtain by the end of the year.

When we consider loans requested in order to enroll in a private higher education institution, we find that a small fraction is however found in public universities according to the Hemis database. This is a case for 18% with a loan and 11% without a loan. Here again, it is not unlikely that a few people changed their plans, but it doesn't seem to be in reaction to a loan refusal: in fact this 7 point difference does not survive a causal estimation (see Figure 6, later). Also, looking at courses and credits, conditional on studying in a public university, those students do not appear different from the rest of the enrolled population.

4 Results

4.1 Empirical strategy

The main objective of this paper is to estimate the causal impact on enrollment in higher education of being granted a loan by Eduloan. With no loss of generality, we consider the following model, estimated over a sample of applicants:

$$Y = \alpha + \beta L + \varepsilon$$

where Y is a dummy for enrollment and L a dummy for loan. α and β are parameters to be estimated and ε is a residual that contains unobserved determinants of enrollment other than Eduloan loan. Because ε may be correlated with L, simple correlation between enrollment and loan does not provide a parameter that has a causal interpretation.

In order to identify a causal impact, we use the regression discontinuity design (see Imbens and Lemieux, 2007, and Lee and Lemieux, 2010, for presentations of this method). We take advantage of the presence of the Empirica score, a credit score E, that strongly influences Eduloan's decision to provide the loan. There is a threshold E_0 that determines eligibility: in principle, Eduloan agents are not to grant a loan if the borrower value of E is below E_0, although there are exceptions.

Figure 1 shows the probability of obtaining a loan, as a function of the threshold (normalized to zero), for loan requests in the Hemis perimeter (i.e. for a public university) for years 2004-2007.[6] Each point represents the proportion of applicants that received a loan among individuals with values of E in a small range. In this graph and the following, we restrict the sample to a neighborhood of plus or minus 100 points around E_0 (total range is about 400 points, but the information is very noisy at large values). On the left of E_0, the probability to obtain a loan is very small, although not strictly zero. Probability to obtain a loan is increasing with the score when the Empirica gets closer to the threshold. There is a very strong discontinuity past the threshold: the probability to obtain a loan jumps from about 10% to about 50%. It then increases smoothly.

The discontinuous relation between L and E at E_0 identifies the causal impact of loan on enrollment if all other determinants (ε in the above statistical model) vary continuously with E, at least in the neighborhood of E_0. Indeed, individuals very close to the threshold have very different proportions of loan access but are otherwise extremely similar. As Lee and Lemieux (2010) convincingly argue, this strategy is in essence very similar to randomization, to the extent that individuals happen to have a few more points in E only by mere chance. This is very arguable in the case of the Empirica, because it is

[6]The value of E_0 remained constant over that period.

based on an unknown algorithm that depends on a number of variables.

This strategy has several limitations. First, identification is local: strictly speaking, it is relevant only for the population close to the threshold. In practice, we will see that, in our data, the population is fairly concentrated around E_0, so that we estimate a parameter that is valid for most of our sample. Second, as shown by Hahn et al. (2001), if treatment effect is heterogenous and if loan access is correlated with loan impact, then the estimated parameter is a local average treatment effect (LATE) in the sense of Imbens and Angrist (1994). [7] In our context, it is not clear that this is a strong limitation because there is no reason that Eduloan agents grant the loan in consideration of the chances to effectively enroll. Indeed, the loan is guaranteed by the customer current income, not on future income that would depend on graduation. Therefore correlation between impact and loan access is not particularly expected.

In order to proceed with estimation, first consider the first-step model that describes the discontinuous relation between loan access and the Empirica score:

$$L = g(E) + \delta D + u \tag{1}$$

where $D = 1$ if $(E \geq E_0)$, $g(E)$ is a continuous functions of E (at least in the neighborhood of E_0), and δ measures the discontinuity jump. This simply fits the data in Figure 1. We can either estimate it on a large range of values of E and use flexible forms for g, or restrict the sample to a neighborhood of E_0 and estimate local linear regressions, that approximate the function as linear. In both cases, specifications allow the function $g(E)$ to be different on the right and on the left of the discontinuity. Similarly, the structural equation can be written as:

$$Y = f(E) + \beta L + \varepsilon' \tag{2}$$

where ε has been expressed as a continuous functions of E. Conditional on E, D

[7] It identifies an average of the causal loan effect on the population who would not have access to loan on the left of E_0 and would get a loan when on the right of E_0.

is a valid instrument for L, so that this model can be estimated by instrumental variable. Here again, $f(E)$ can be allowed to have different shapes on the right and on the left of the discontinuity, and the model can be estimated on a large range or by local linear (instrumental) regression. In the latter case, bias is minimized when the sample is strongly restricted to the neighborhood of E_0 but precision is increased as the sample gets larger. Imbens and Lemieux (2007) suggest a cross-validation procedure in order to select the optimal bandwidth in terms of mean squared error. Depending on the specification, we find optimal bandwidths of +/- 65 or +/- 125 Empirica points around the threshold[8]: these are quite large bandwidth, and it reflects the fact that the linear approximation is adequate on a large range. Nonetheless, we always present regressions for the full sample, using linear or quadratic functions for g or f, with different slope on either part of the discontinuity, and local linear regression for different bandwidth, including the optimal one.

Table 2 present the estimation of equation 1: the increase in the proportion treated due to the discontinuity is evaluated between 0.32 and 0.42 depending on the specification, always very significant. At the optimal bandwidth of +/- 65 points, the effect is 0.39 and it is only slightly lower (0.36) using the full sample with quadratic functions. This ensures that the instrument will have identifying power.

We can also check that E_0 is not a threshold for other variables than loan. Table 3 shows that there is no discontinuous change of the wage of the borrower, of the choice of a public or a private institution, and of the amount of loan requested. This confirms that borrowers are not aware of their Empirica score or of the threshold, so that they do not ask for larger loans when they know that their chances to be accepted are strong. Finally, Figure 2 plots the density of observations around E_0. First, there is no evidence of bunching to the right of the threshold, which would happen if individuals could manipulate their Empirica at the margin. Second, we can see that observations are concentrated

[8]Optimal bandwidth is mentioned in table notes.

around the threshold, so that, as mentioned earlier, local identification still involves a large fraction of our sample.

4.2 Baseline results: impact of loans on enrollment

Table 4 and Figure 3 show the reduced form relation between enrollment and the Empirica score. The probability to be enrolled at a public university, for individuals who asked a loan in order to pay fees for such university, increases precisely at the threshold E_0. This should not happen if loan wasn't a causal determinant of enrollment, unless there are other determinants of enrollment that change discontinuously also at E_0, something that we argued could be excluded in this environment. The effect is strong and very significantly estimated at 9 to 10 percentage points. Given that the threshold value is normalized to zero, the enrollment rate just at the left of the discontinuity is directly given by the constant: therefore, this reduced form effect is to increase enrollment rates from about 50% to about 60%.

Table 5 presents estimates of equation 2. Ordinary least square estimation indicates that loan obtention increases enrollment by 20 percentage points. Instrumental variable estimation, using the discontinuity as an instrument, raises this effect to about 22 to 25 points. A stronger effect is found for the small +/- 20 bandwidth, but this is the exception and this range is not the optimal one. As a result, we can claim that providing a loan to this population causally increases the chances that it will enroll in higher education from a level of 50% to 73%, at least for individuals close to the Empirica threshold. As expected, the results hardly change if we add control variables such as age, gender, required loan amount or monthly wage, because the instrument is not correlated to those variables. Including them does not systematically improve the precision of the estimation, so we present the simple regressions that are more transparent.

OLS estimation appears to be biased (precision is sufficient for a Hausman test to reject equality of the OLS and IV parameters), but the size of the bias is small. It implies that characteristics observed by Eduloan that determine loan acceptance are marginal determinants of the individual decision to enroll in this sample.

We do not find any significant difference when measuring loan impact separately for men and women. It also does not seem to make a difference whether the borrower is the student himself or a relative. However, as shown in Table 6, the impact of the loan is different among the richest and the poorest borrowers. We do not have a lot of statistical power when it comes to splitting the sample, but we can distinguish between the lowest wage quartile and the rest of the sample (higher panel) or between above and below the median (lower panel). Loan impact is about twice as large for the lowest quartile and about 70% higher when we compare samples across the median. Although the first comparison is only significant at the 10% level, and the second comparison is not significant at all, this is indicative of a plausible fact: that credit constraint is stronger for less wealthy families and that less financing alternatives exist at the bottom of the income distribution. One possibility is that commercial banks may be willing to grant loans to some of the richest individuals in our sample, thereby diminishing the impact of Eduloan activities on this specific population.

Two other outcome variables are shown in Table 7. The *number of courses registered for* takes value zero for the non enrolled and whatever positive values for the enrolled, and similarly for *credits completed*. Because they enroll more frequently, applicants who get a loan tend to register for more courses on average (1.5 more courses, a 44% increase at the optimal bandwidth specification) and complete more credits than those who are rejected (around 8 percentage points, a 39% increase). In South Africa, one year of higher education consists in 1.0 credit, so that a typical academic year is made of 10 courses, each one worth 0.1 credit: our descriptive statistics recall the low completion rate of students, whether they get a loan or not. As discussed below, we cannot identify the impact of having a loan on educational outcomes *conditional on enroll-*

ment[9] However, we were able to show that increased enrollment resulting from loan access does translate into increased registration and credit completion, which is important from a policy point of view.

5 Robustness

5.1 The 2008 credit crunch

In 2008, the financial crisis induced a restriction in credit that impacted Edu-loan among other financial institutions. As a result, fewer loans were granted that year, especially to people above the Empirica threshold, as illustrated in Figure 4 for years 2007 and 2008. We can thus compare individuals on the right of the Empirica score before and after 2008 and use this as a different identifying information to check the robustness of our initial results.

Figure 5 shows the reduced form relationship between enrollment rates and the Empirica threshold: the discontinuity that is apparent in 2007 disappears in 2008, and this mirrors perfectly the structure of loan access in Figure 4. We can fit this data with a model that interacts functions $f(S)$ in equation 2 with years:

$$Y = f_{2007}(S) + f_{2008}(S) + \theta D + \beta L + \varepsilon'$$

In this regression we can allow D to be present in the regression because L is now instrumented by the interaction between D and year 2007: we thus use a different identification restriction. As a result, this also gives an opportunity to check that being on the right-hand side of the discontinuity has no impact on enrollment when it has no impact on loans: we expect $\theta = 0$.

[9]If we compare individuals with and without a loan among the enrolled, we mix two effects. One is that loan induces a different performance of *ex ante* similar people in the two groups, the other is that loan induces enrollment of additional people, and those people may be different in terms of academic capacity or motivation. This is the usual selectivity issue, as faced by Canton and Blom (2004) for instance. Because we do not have an exogenous determinant of selection that would not have a direct influence on performance, we cannot control for selection without making arbitrary parametric assumptions. Bounds analysis only generates uninformative bounds here.

Table 8 presents this estimation for 2007 and 2008.[10] Although they are based on a different type of information, coefficients on loans are only slightly smaller than in the baseline estimation, but they are very comparable and significant. Also, the coefficient on D is small and non significant, which confirms our baseline identification hypothesis.

5.2 Sample variants

As mentioned earlier, the sample used until now has been restricted to loans requested to pay public university fees (Hemis perimeter), but only when information on the kind of university was actually available. There are 2,664 observations for which either the field was not filled, or for which the abbreviation or acronym used did not refer to an institution we could clearly identify. This sample may contain a number of loans in the Hemis perimeter, and the corresponding population may be specific. As a robustness check, we would like to include this population. However, this implies including also an unknown proportion of loans requested for private institutions.

The appendix shows that, if we do so, and estimate the regression discontinuity model using both the known public and unknown samples, we obtain a lower bound to the true parameter. This is true provided that loan access has no causal effect on enrollment in a public university for those who wished to enter a private institution. This is expected given the fact that fee payment is delivered directly by Eduloan. It is confirmed by Figure 6: this figure uses the sample of loan requests known to be for a private institution (2,473 observations) and plots the reduced form of enrollment in a public university as a function of the Empirica discontinuity. There is no evidence that loan status has any impact on enrollment in a public university.[11]

In this context, it is intuitive that pooling public and non-public loan demands will provide an average of: (1) the true effect on Hemis perimeter de-

[10] In 2008, applications past April are excluded because the Empirica threshold E_0 was increased.

[11] Remember that a small share of individuals who asked a loan for a private university end up enrolled in a public university. Figure 6 shows that this is unrelated to loan status.

mands; and (2) a zero effect, thus a lower bound to the true effect. As detailed in the appendix, this argument is complicated by the fact that the two subsamples may have different discontinuity impacts in the first-stage regression, but we show that the lower bound rule remains. Results are presented in Table 9 and they show significant effects, to the order of 0.18, to be compared to our baseline estimates of about 0.23 (Table 5). We are thus confident with the presence of an impact and its order of magnitude.

A second restriction to our baseline sample has been to exclude observations with loan requests placed in November or December, because we are unsure whether they refer to the current year or to the next year. The sample change is rather marginal, as the number of observations is increased by only 12% if we keep late requests. With such data, we expect some enrollment measures to correspond to the wrong year. According to the same argument as above, the impact has to be zero for a (small and unidentified) share of the sample, because the wrong year outcome variable will not be sensitive to loan access. Including November and December requests, we thus estimate a lower bound. Table 10 shows that coefficients are only slightly lower than our baseline estimates.

To sum up, data limitations imply that, strictly speaking, our baseline estimation may have external validity limitation, even if we restrict our universe to loan demands to Eduloan in order to enroll in public universities. When we enlarge the sample, we can only estimate bounds to the parameter of interest. Yet, such bounds do confirm the order of magnitude of the effects and they are not significantly different from our baseline point estimates.

6 Discussion

6.1 Enrollment in the private sector

We have shown that, when an individual plans to enter a public university and asks Eduloan a short-term loan to pay the fees, he or she is more likely to enroll in a public university on that same year when the loan was granted. We cannot strictly exclude that an individual whose demand is turned down decides to enroll in the private sector instead, because our data contains no information on private enrollment.[12] To the extent that our major questioning is on the existence of a liquidity constraint and the estimation of how many individuals are constrained in a population, our conclusion is robust: a large number of individuals who had an explicit plan to enter some kind of university had to change this plan one way or another because they did not get short-term credit to afford the fees in that university.

It is more arguable whether this liquidity constraint results in an equivalent decrease in the number of individuals that actually enter higher education. If private institutions are less expensive than public universities[13], it could be rational for some individuals to turn to a private institution when they are declined a loan by Eduloan, provided cost is sufficiently low to avoid the liquidity constraint, and quality is sufficiently high to make this choice a second best. If such behaviour (unobserved by us) was present, this would reduce the loan impact *in terms of overall enrollment in higher education*. We do know from Figure 6 that the reverse does not hold true: individuals who apply to a private institution, and are not granted a loan for that, do not turn more often to a public university. But this could be because this is just a more expensive option.

Yet, we have a way to check whether individuals denied a loan by Eduloan tend to chose to enter a less costly university instead. South Africa has a famous distance learning institution, which was open to black and couloured persons under apartheid: the University of South Africa (UNISA). In our data, 31% of all loan demands for a public university (Hemis perimeter) are made for UNISA. Its lower cost is reflected in the amount of required loans: the average loan request is ZAR 8,051 for other public universities but only ZAR 3,885

[12]As a matter of fact, there are a few individuals who have filed loan requests for public and for private institutions. When this is the case, the year-loan demand observation has been classified as private, in order to remain on the safe side.

[13]Anecdotal evidence tends to indicate this is the case, algthough there is substantial heterogeneity.

for UNISA. Table 11 uses individuals that asked a loan for a public university *other than* UNISA. It checks whether the individuals who, among them, were declined the loan, are eventually found registered at UNISA. To do so, we simply use the same regression discontinuity design as before to estimate the causal effect of a loan on this new outcome ("being registered at UNISA"). We find no evidence of such a behavior.

If shifting to a less costly institution was optimal for many individuals when a loan for a public university is declined, then we would expect that at least some of them would choose to shift to UNISA while some others would enter a private university. As we find no evidence of the former (in spite of the fact that UNISA is a well known and popular institution), we do not expect that the latter should be a major source of bias to the enrollment impact of loans.

6.2 External validity

We have already mentioned limitation to external validity due to the specific population of Eduloan applicant. On the one hand, we do not include people for whom higher education is not an option anyway, on the other hand, among the rest of them, we probably do not observe the poorest.

Eduloan provides its clients with short-term loans with a limited grace period (most of the time 2 or 3 months). This implies that Eduloan only alleviates short-term constraints, that make for only a part of financial credit constraints. Stronger solvency issues, that can only be solved through an increase in future revenues and longer grace periods, are not identified in our study, since students who cannot repay at least part of their loan while studying are not granted a loan. The fact that even a simple smoothing payment mechanism has a very significant impact on university enrollment suggests that the credit constraint must be very strong in South Africa. Eduloan, as simple a mechanism as it may seem, is nonetheless a unique system in the developing world, with no known equivalent in Africa for example.

This general result is all the more striking that South Africa is a highly financialized emerging country. Its credit-to-GDP ratio amounts to 88% in 2009, much higher than that of Burkina Faso (15%), Cameroon (23%), Nigeria (26%), Ghana (32%) or Kenya (35%). It indicates a level of financial development close to that of other emerging countries such as Vietnam or Thailand (between 90% and 100% according to the IMF). Since the level of financial development is correlated with GDP per capita, and South Africa is a relatively rich country in the developing world, we would expect credit constraints to be even more significant in most of the world. Therefore, it is most likely that many students in low-income or middle-income countries are also strongly affected by credit constraints, limiting their ability to achieve the studies their talent would allow. In that context, the development of education loan mechanisms seem suitable.

6.3 Cost-efficiency

Are student loan mechanisms costly to support? As mentioned above, although Eduloan is a private company, it has several partnerships with international donors. At least three development finance institutions partnered with Eduloan but none ever subsidized it. As an example, the French Development Agency guaranteed 50% of the amount of loans taken by Eduloan from a South African bank, against an annual fee.

Such a risk-sharing agreement between a development agency and a local bank has no direct cost for the donor unless the eventual borrower goes bankrupt or asks for some form of debt cancelation or restructuring. If there was a market for such guarantees, the annual cost of this "development project" would be:

$$c = A(p^* - p)$$

where c is the cost for the development agency, A the amount guaranteed (commonly 50% of the total loan taken by the borrower), p^* the annual market price of such a guarantee (a form of interest rate) and p the annual price actually paid by the local institution. While p^* is not an observable parameter, we can figure out orders of magnitude. Grossly, it must be a function of the borrower bankruptcy risk and the return expected by the local bank (and the donor agency) on its off-balance commitments. If we assume that Eduloan is comparable to the average South African SMEs, market rates for SMEs give an indirect indication of possible p^* values.

In South Africa, the Central Bank sets a repurchase rate (also called "Repo", comparable to US Federal Funds rate or the European Central Bank refinancing rate) and a Prime overdraft rate (also called "Prime"). The Prime rate is 3 to 4% higher than the Repo and is a reference rate for households or SMEs, that generally borrow money paying the prime rate plus one or two percentage points. Since commercial banks then have a 5% margin over the refinancing rate when they lend to customers, the usual 1% bank operating costs lead to a market price of the risk coverage of 4%.

If p^* was equal to 4% in our example, p would have to be smaller than 4%, at least by 0.5%[14] and most likely bigger than 1% (because Development banks also have operating costs). We can therefore assume that p stand somewhere between 1% and 3.5%, so that c should be in the $[0.005A \; ; \; 0.03A]$ interval. This seems a modest cost on public money (roughly 2% of loan amounts according to this estimation) for a "program" that increases by 50% university enrollment in a population.

7 Conclusion

Having access to customers data from a private credit company (Eduloan) enables us to provide direct evidence on the impact of credit constraint on higher education enrollment in South Africa, whereas most of the related literature relies on indirect or subjective evidence. Eduloan uses a threshold to grant its loans, which allows us to implement a robust identification strategy based

[14] Otherwise there would not be any real interest for the local bank and no development finance institution would be necessary.

on a regression-discontinuity design. The causal impact of access to credit is estimated for a relevant population, that is the individuals willing to borrow and enroll at a South African university.

We show that our sample is strongly constrained by liquidity and that obtaining a loan raised the probability to enroll by about 23 percentage points between 2004 and 2007, a 50% increase, and raised the number of academic credits completed by borrowers by roughly 40%. We also find that effects are stronger for the poorer part of our sample, which confirms the notion that such constraint should be more binding for that population. Therefore, although South Africa is a highly financialized country, liquidity constraint matters for the access to higher education. It may be even stronger in many other low- and middle-income countries where financial markets are even more incomplete.

One important difference of our findings with respect to mostly US based evidence, appart from methodology, is that either credit markets for human capital investment are more present (as analyzed by Lochner and Monge-Naranjo, 2011) or the large range of subsidies to education that exist compensate for credit market constraints more than in the developing world. To that extent, the mixed evidence from most of the literature is a poor guide for higher education policy in the developing world and this paper is one of the very few so far to fill the gap.

On the policy side, our findings tend to support State or Donor sponsored loan schemes, at least in developing countries, as they are likely to offer both efficiency and equity benefits. Several such schemes do exist already but, to our knowledge, they have not been evaluated in terms of impact. This would be desirable in order to confirm the generality of our conclusions.

References

[1] Acemoglu D. & Pischke J-S. (2001), "Changes in the wage structure, family income, and children's education", *European Economic Review*, Vol.

45, pp. 890-904.

[2] Attanasio O.P. & Kaufmann K.M. (2009), "School Choices, Subjective Expectations, and Credit Constraints", NBER Working Paper No. 15087.

[3] Branson N., Leibbrandt M. and Zuze T. L., (2009), "What are the returns of tertiary education and who benefits?" in *Responding to the Educational Needs of Post-school Youth*, N. Cloete (Ed.), Centre for Higher Education Transformation.

[4] Brown M., Scholz J.K. & Seshadri A. (2009), "A new test of borrowing constraints for education ", NBER Working Paper No. 14879.

[5] Cameron S.V. & Heckman J.J. (2001), "The Dynamics of Educational Attainment for Black, Hispanic, and White Males", *Journal of Political Economy*, Vol. 109, No. 3, pp. 455-499.

[6] Cameron S.V. & Taber C. (2004), "Estimation of Educational Borrowing Constraints Using Returns to Schooling", *Journal of Political Economy*, Vol. 112, No. 1, pp. 132-182.

[7] Canton E. and Blom A. (2004), "Do student loans improve accessibility to higher education and student performance. An impact study of the SOFES program in Mexico", CPB Discussion paper 33, The Hague.

[8] Card D. (2001), "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems", *Econometrica*, Vol. 69, No. 5 (Sep., 2001), pp. 1127-1160.

[9] Carneiro P. & Heckman J.J. (2002), "The Evidence on Credit Constraints in Post-Secondary Schooling", *The Economic Journal*, Vol. 112, No. 482, pp. 705-734.

[10] Department of Education (2009), Education Statistics in South Africa 2007, Department of Education, South Africa.

[11] Hahn J., Todd P. & Van der Klaauw, (2001), "Identification and Estimation of Treatment Effects with a Regression Discontinuity Design", *Econometrica*, Vol. 69, No. 1, pp. 201-210.

[12] Imbens G. & Angrist J. (1994),"Identification and estimation of local average treatment effects", *Econometrica*, vol. 62, pp. 467-475.

[13] Imbens G., & Lemieux T. (2007), "Regression Discontinuity Designs: A Guide to Practice", *Journal of Econometrics*, Vol. 142, No. 2, pp. 807-828.

[14] Keane P.K. & Wolpin K.I. (2001), "The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment", *International Economic Review*, Vol. 42, No. 4, pp. 1051-1103.

[15] Keswell M. and Poswell L. (2004), "Returns to Education in South Africa: A retrospective Sensitivity Analysis of The Available Evidence", *South African Journal of Economics*, Vol. 72, No. 4, pp. 834-860.

[16] Lang K. (1993), "Ability Bias, Discount Rate Bias and the Returns to Education", unpublished discussion paper, Boston University.

[17] Lang K. and Ruud P. (1986), "Returns to Schooling, Implicit Discount Rates and Black-White Wage Differentials", *The Review of Economics and Statistics*, Vol. 68, No. 1, pp. 41-47.

[18] Lee D. & Lemieux T. (2010), "Regression Discontinuity Designs in Economics", *Journal of Economic Literature*, Vol. 48, No. 2, pp. 281-355.

[19] Lochner L. & Monge-Naranjo A. (2011), "The Nature of Credit Constraints and Human Capital", *American Economic Review*, forthcoming.

[20] Maurin E. (2002), "The Impact of Parental Income on Early Schooling Transitions : A Re-examination Using Data over Three Generations", *Journal of Public Economics*, Vol. 85, No 3, pp. 301-332.

[21] Social Surveys (2009), "Access to public higher education in South Africa: Financial constraints and consequences", mimeo, Social Surveys, South Africa.

[22] Statistics South Africa (2006), "Quarterly Employment Statistics, June 2006, September 2006, December 2006".

[23] Stinebrickner R. & Stinebrickner T. (2008), "The Effect of Credit Constraints on the College Drop-Out Decision: A Direct Approach Using a New Panel Study", *The American Economic Review*, Vol. 98, No. 5, pp. 2163-2184.

[24] Stumpf R. et al. (2008), *Tuition Fees: Higher Education Institutions in South Africa*, HESA Report, Higher Education South Africa.

Appendix: a lower bound to the estimator when we mix Hemis and non-Hemis loan requests.

We are interested in the parameter $E[Y(1) - Y(0)|E = E_0, H = 1]$ where $Y(1)$ is counterfactual enrollment when a loan is granted and $Y(0)$ when it is not. E is the Empirica score, E_0 being the identifying threshold, and $H = 1$ if the individual asked a loan for a Hemis (i.e. public) institution and $H = 0$ otherwise. The parameter is defined for the Hemis population and the problem stems from the fact that we do not observe H (or not fully so). We use the notation $E^+[.|E = E_0] = \lim_{E\to E_0^+} E(.|E)$ for the right-hand-side limit to the threshold and similarly with minus for the left-hand-side. Following Hahn, Todd & Van der Klaw (2001), let us think in terms of the Wald estimator. We can compute empirically:

$$W = \frac{E^+[Y|E_0] - E^-[Y|E_0]}{E^+[L|E_0] - E^-[L|E_0]}$$

where Y is observed outcome and L is observed loan status (obtained or not). The Wald estimator W is the parameter we compute using any of the 2SLS methods devised in the text when we pool Hemis and non-Hemis loan demands.

For the students who applied to a private university, the public enrollment variable is always equal to zero in our data, that is formally: $Y(0) = Y(1) = 0$ if $H = 0$. Then:

$$E(Y|E) = P(H = 1|E)E[Y|E, H = 1] + 0$$

by construction and, $P(H = 1|E)$ being continuous in E_0:

$$E^+[Y|E_0] - E^-[Y|E_0] =$$

$$P(H = 1|E_0)\cdot E[Y(1)-Y(0)|E_0, H = 1]\times(E^+[L|E_0, H = 1]-E^-[L|E_0, H = 1])$$

In addition we have:

$$E^+[L|E_0] - E^-[L|E_0] =$$

$$P(H=1|E_0) \times \left(E^+[L|E_0, H=1] - E^-[L|E_0, H=1]\right) +$$
$$(1-P(H=1|E_0)) \times \left(E^+[L|E_0, H=0] - E^-[L|E_0, H=0]\right).$$

Replacing in the first equation, it is straightforward to show that:

$$W = E[y(1) - y(0)|E_0, H=1] \times$$
$$\left[1 + \frac{1-P(H=1|E_0)}{P(H=1|E_0)} \frac{E^+[L|E_0, H=0] - E^-[L|E_0, H=0]}{E^+[L|E_0, H=1)] - E^-[L|E_0, H=1)]}\right]^{-1}$$

The term within bracket is clearly positive and higher than 1 so that we can write that

$$W =< E[y_1 - y_0|E_0, H=1]$$

which in turn means that W estimates a lower bound to the parameter of interest.

Figure 1: Share of loans granted as a function of Empirica score (2004-2007)

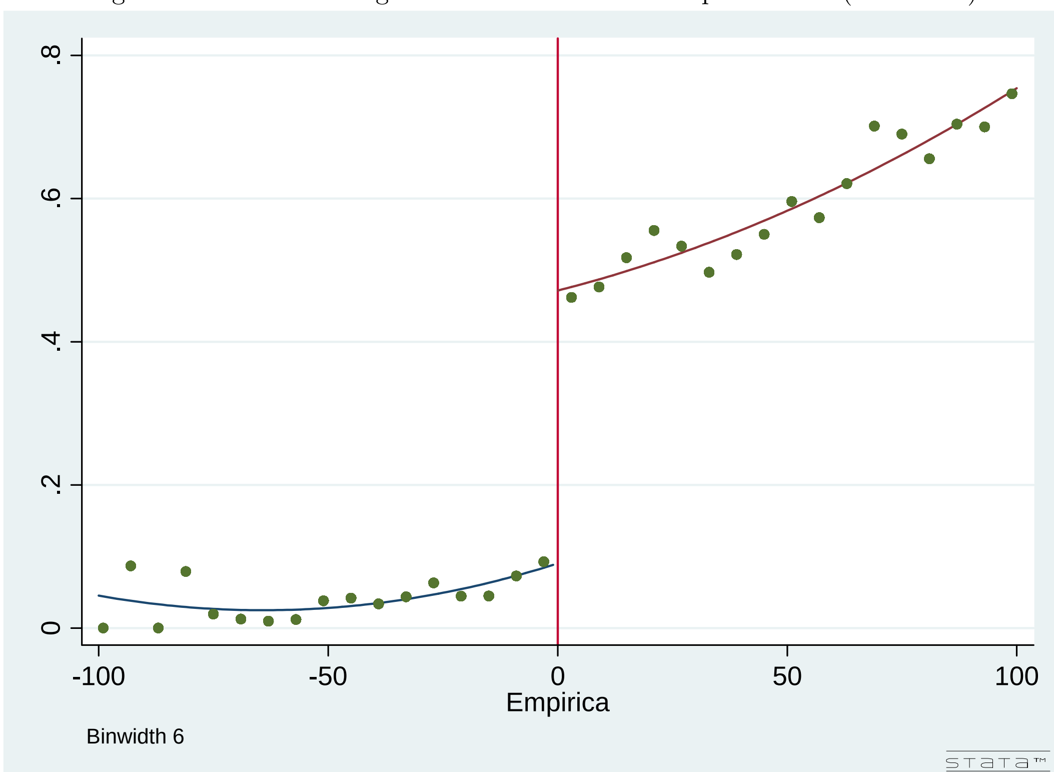

Figure 2: Density of the Empirica score (2004-2007)

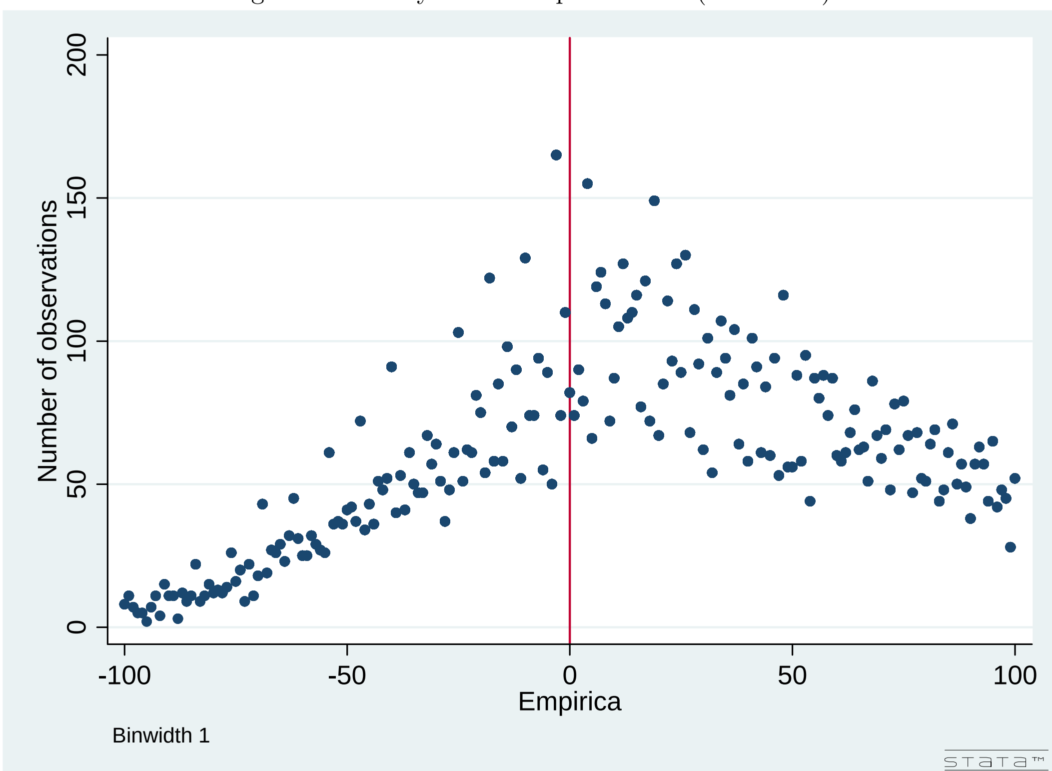

Figure 3: Proportion of university enrollment as a function of Empirica score (2004-2007)

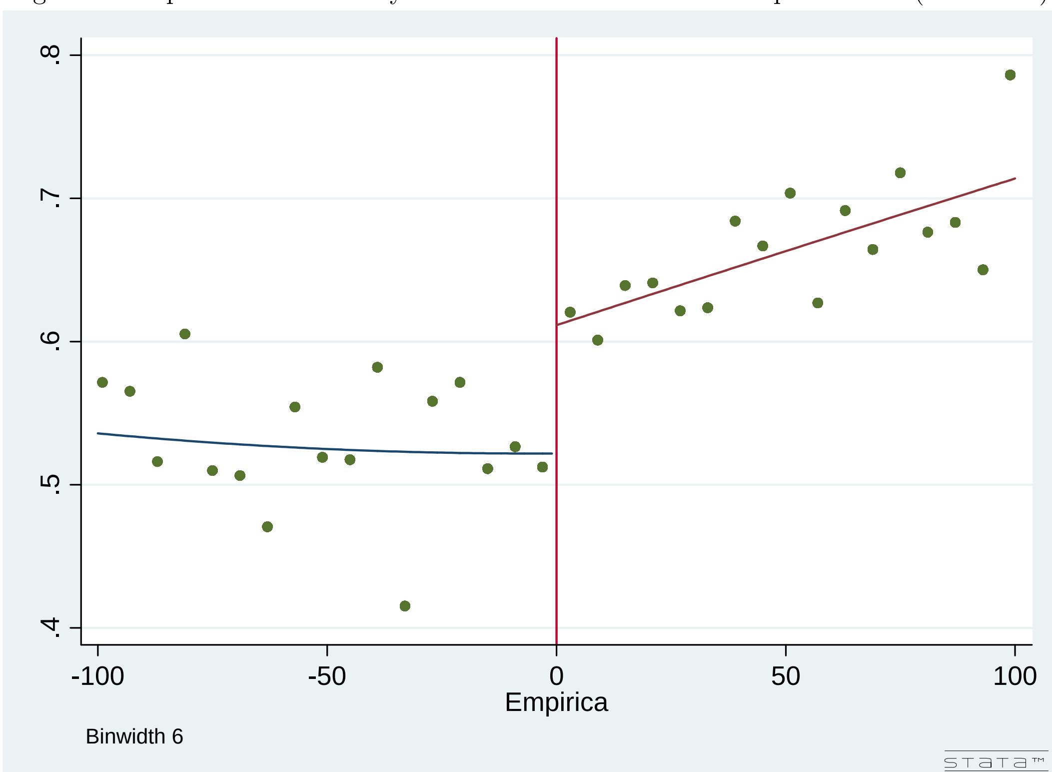

Figure 4: Share of loans granted as a function of Empirica score, 2007 and 2008

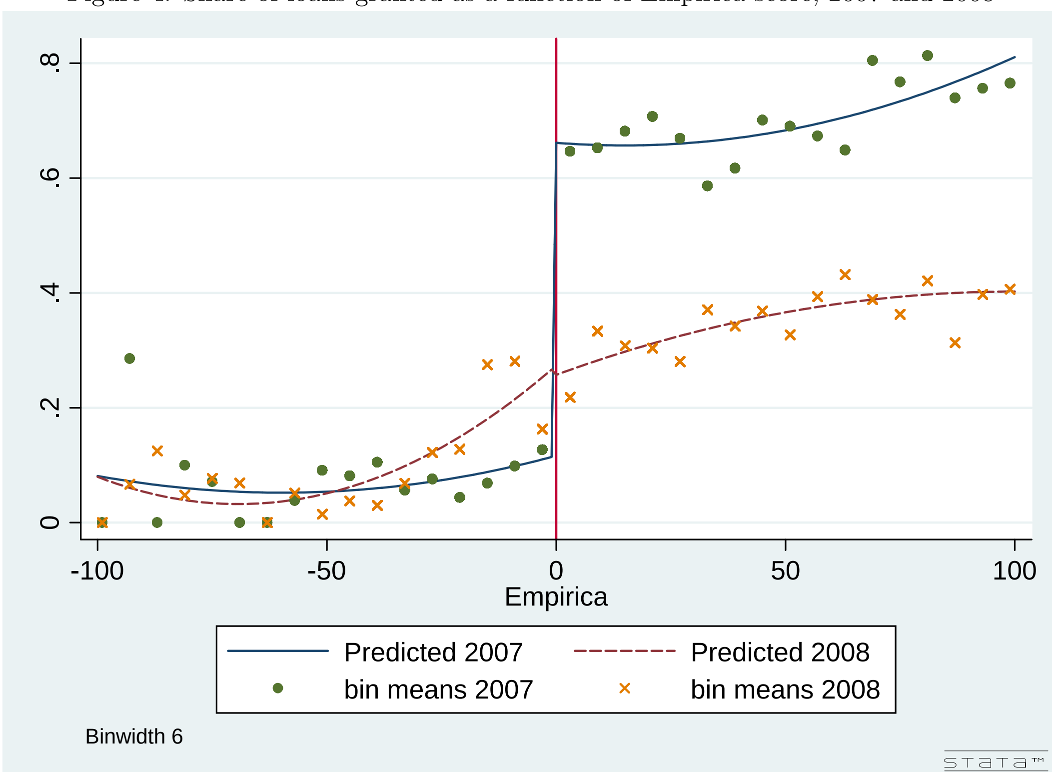

Figure 5: Proportion of university enrollment as a function of Empirica score, 2007 and 2008

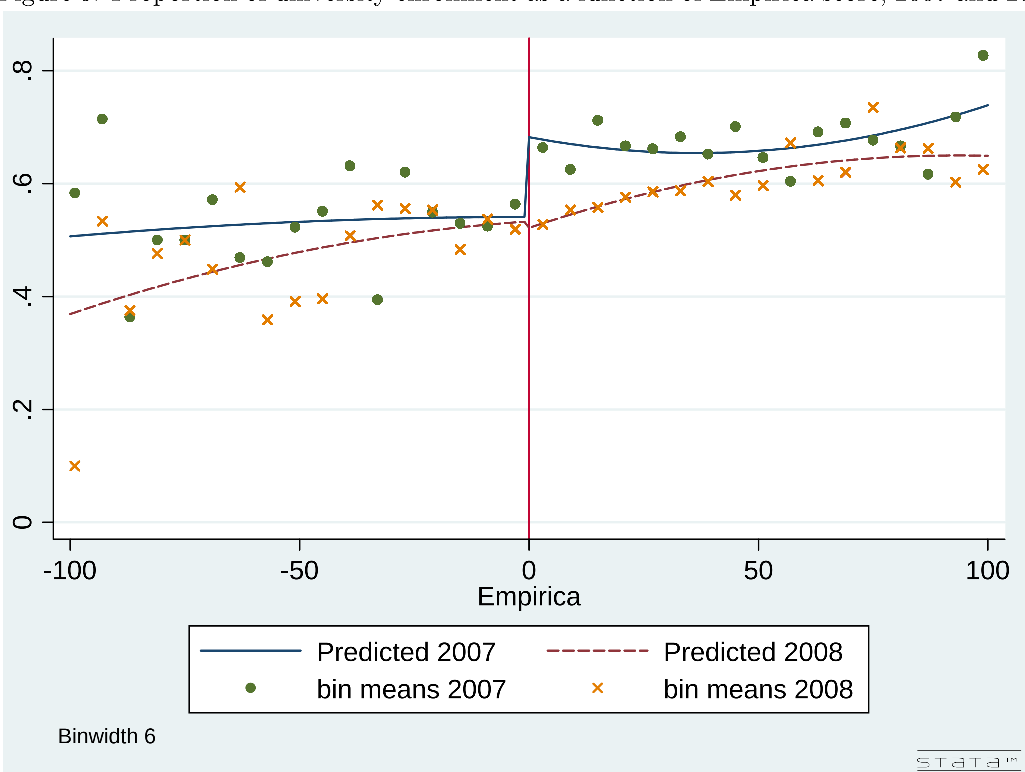

Figure 6: Proportion of Public University enrollment for individuals who requested a loan for a Private University, as a function of Empirica score (2004-2007)

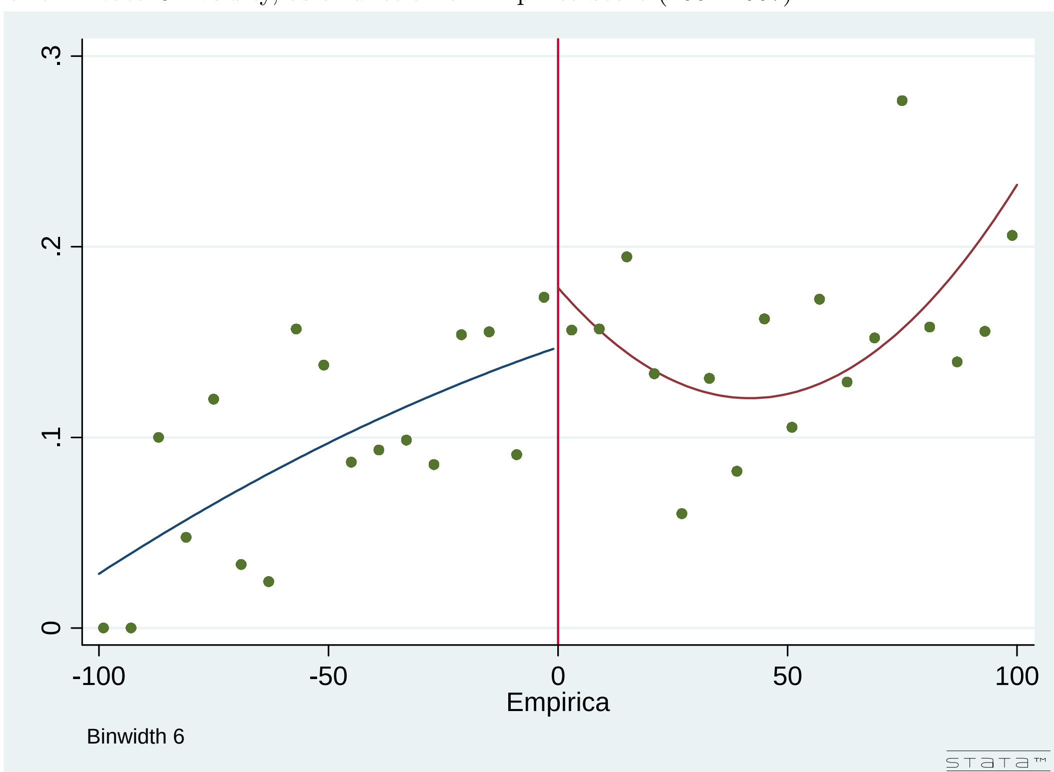

Table 1: Descriptive statistics on loan demands, 2004-2007

	Loan requested for a public institution (Hemis perimeter)			
	No loan obtained		Loan obtained	
	mean	s.e.	mean	s.e.
Male	0.46	0.50	0.48	0.50
Age	27.83	8.41	27.57	7.86
Monthly wage	6 420	5 018	7 521	7 365
Missing wage information	0.15	0.36	0.00	0.00
Requested loan/monthly wage	1.53	1.65	1.04	0.80
Missing requested loan value	0.03	0.18	0.00	0.06
Student is the borrower	0.49	0.50	0.49	0.50
Enrolment in public University	0.53	0.50	0.75	0.43
Credits completed (if enrolled)	0.46	0.37	0.44	0.37
Number of courses registered (if enrolled)	7.14	4.37	6.82	4.13
# observations	5 165		4 814	

	Loan requested for a private institution			
	No loan obtained		Loan obtained	
	mean	s.e.	mean	s.e.
Male	0.46	0.50	0.50	0.50
Age	26.46	9.30	26.18	8.55
Monthly wage	5 918	4 339	6 736	4 929
Missing wage information	0.15	0.35	0.00	0.00
Requested loan/wage	2.18	2.23	1.44	1.26
Missing requested loan value	0.07	0.26	0.06	0.24
Student is the borrower	0.35	0.48	0.36	0.48
Enrolment in public University	0.11	0.31	0.18	0.38
Credits completed (if enrolled)	0.44	0.35	0.44	0.35
Number of courses registered (if enrolled)	7.41	4.36	7.41	4.36
# observations	1 707		766	

	Institution unreported or ambigous			
	No loan obtained		Loan obtained	
	mean	s.e.	mean	s.e.
Male	0.48	0.50	0.48	0.50
Age	27.24	8.42	27.27	8.43
Monthly wage	5 942	4 695	6 880	5 008
Missing wage information	0.11	0.32	0.00	0.00
Requested loan/wage	1.58	1.68	1.10	0.85
Missing requested loan value	0.14	0.35	0.33	0.47
Student is the borrower	0.44	0.50	0.43	0.49
Enrolment in public University	0.51	0.50	0.74	0.44
Credits completed (if enrolled)	0.47	0.37	0.47	0.36
Number of courses registered (if enrolled)	7.42	4.33	7.56	4.09
# observations	1 896		768	

Note: The unit of observation is loan demand per year. Loans requested in November/December are excluded as in all baseline estimations. When several requests have been sent for a given student the same year, we use the average requested loan.

Table 2: Loan granted as a function of Empirica score (Hemis perimeter, 2004-2007)

	Total sample		Local linear regression for various bandwidth			
			+/-125	+/- 100	+/- 65	+/-20
Above discontinuity point	0.3589	0.4247	0.3884	0.3790	0.3877	0.3284
	0.0177	*0.0126*	*0.0141*	*0.0152*	*0.0181*	*0.0320*
Empirica	0.0016	0.0007	0.0007	0.0008	0.0011	0.0034
	0.0005	*0.0002*	*0.0002*	*0.0002*	*0.0003*	*0.0013*
Empirica x above	0.0021	0.0010	0.0019	0.0020	0.0010	0.0035
	0.0006	*0.0002*	*0.0002*	*0.0003*	*0.0005*	*0.0026*
Empirica sq	0.0094					
	0.0039					
Empirica sq x above	-0.0208					
	0.0044					
Intercept	0.0854	0.0727	0.0731	0.0748	0.0816	0.1011
	0.0110	*0.0079*	*0.0082*	*0.0085*	*0.0096*	*0.0172*
# observations	9 979	9 979	8 531	7 717	6 011	2 340

Sample: loans requested for registration in a public University ("Hemis perimeter"). The unit of observation is loan demand per year. Loans requested in November/December are excluded as in all baseline estimations. Explained variable is a dummy that takes value 1 when a loan has been granted. Ordinary least squares. Robust standard errors in italics. Bandwidth of +/-65 around discontinuity point is the optimal bandwidth according to the cross-validation criteria.

Table 3: Predetermined variables as a function of Empirica score (2004-2007)

	Log wage	Loan requested for public university (as opposed to private)	Log requested loan amount
Above discontinuity point	-0.0099	-0.0002	0.0012
	0.0260	*0.0191*	*0.0354*
Empirica	0.0016	0.0025	-0.0002
	0.0010	*0.0008*	*0.0014*
Empirica x above	0.0193	0.0164	-0.0027
	0.0091	*0.0082*	*0.0141*
Empirica sq	-0.0012	-0.0016	0.0000
	0.0011	*0.0009*	*0.0015*
Empirica sq x above	-0.0173	-0.0189	0.0092
	0.0094	*0.0083*	*0.0144*
Intercept	8.5896	0.7837	8.5543
	0.0216	*0.0163*	*0.0299*
# observations	13 887	12 452	14 242

Sample: loans requested with non-missing values for the relevant variable. The unit of observation is loan demand per year. Loans requested in November/December are excluded as in all baseline estimations. Ordinary least squares. Robust standard errors in italics. Intercept of log wage and loan amounts not reported for confidentiality reasons.

Table 4: University enrolment as a function of Empirica score (Hemis perimeter, 2004-2007)

	Total sample		Local linear regression for various bandwidth			
			+/-125	+/-100	+/- 65	+/-20
Above discontinuity point	0.0905	0.1017	0.0973	0.0913	0.0873	0.1049
	0.0256	*0.0188*	*0.0200*	*0.0210*	*0.0242*	*0.0409*
Empirica	-0.0004	-0.0001	-0.0002	-0.0001	-0.0001	-0.0014
	0.0011	*0.0004*	*0.0004*	*0.0004*	*0.0006*	*0.0027*
Empirica x above	0.0017	0.0008	0.0011	0.0011	0.0012	0.0025
	0.0011	*0.0004*	*0.0004*	*0.0005*	*0.0007*	*0.0035*
Empirica sq	-0.0031					
	0.0106					
Empirica sq x above	-0.0002					
	0.0108					
Intercept	0.5174	0.5216	0.5177	0.5203	0.5222	0.5042
	0.0217	*0.0163*	*0.0166*	*0.0171*	*0.0191*	*0.0313*
# observations	9 979	9 979	8 531	7 717	6 011	2 340

Sample: loans requested for registration in a public University ("Hemis perimeter"). The unit of observation is loan demand per year. Loans requested in November/December are excluded as in all baseline estimations. Explained variable is a dummy that takes value 1 when a the student is found enrolled at a public University on the same year as the loan request. Ordinary least squares. Robust standard errors in italics. Bandwidth of +/-125 around discontinuity point is the optimal bandwidth according to the cross-validation criteria.

Table 5: University enrolment as a function of loan obtention (Hemis perimeter, 2004-2007)

	Total sample		Total sample		IV for various bandwidth			
	OLS	OLS	IV	IV	+/-125	+/-100	+/- 65	+/-20
Loan	0.2037	0.2050	0.2520	0.2394	0.2506	0.2408	0.2253	0.3194
	0.0108	*0.0105*	*0.0705*	*0.0439*	*0.0508*	*0.0547*	*0.0616*	*0.1232*
Empirica	-0.0002	0.0000	-0.0008	-0.0002	-0.0004	-0.0003	-0.0003	-0.0024
	0.0008	*0.0003*	*0.0012*	*0.0004*	*0.0004*	*0.0005*	*0.0007*	*0.0030*
Empirica x above	0.0009	0.0004	0.0011	0.0006	0.0007	0.0007	0.0010	0.0013
	0.0010	*0.0003*	*0.0011*	*0.0004*	*0.0004*	*0.0005*	*0.0007*	*0.0034*
Empirica sq	-0.0011		-0.0055					
	0.0088		*0.0109*					
Empirica sq x above	v		0.0051					
	0.0085		*0.0117*					
Intercept	0.5122	0.5174	0.4959	0.5042	0.4994	0.5023	0.5038	0.4719
	0.0120	*0.0093*	*0.0270*	*0.0191*	*0.0197*	*0.0205*	*0.0233*	*0.0418*
# observations	9 979	9 979	9 979	9 979	8 531	7 717	6 011	2 340

Sample: loans requested for registration in a public University ("Hemis perimeter"). The unit of observation is loan demand per year. Loans requested in November/December are excluded as in all baseline estimations. Explained variable is a dummy that takes value 1 when a the student is found enrolled at a public University on the same year as the loan request. In IV specification, the excluded instrument is a dummy for "above discontinuity point". Robust standard errors in italics. Bandwidth of +/-65 around discontinuity point is the optimal bandwidth according to the cross-validation criteria.

Table 6: University enrolment as a function of loan obtention (Hemis perimeter, 2004-2007)

Heterogenous effects

	Borrower wage below first quartile (IV)			Borrower wage above first quartile (IV)		
	Total sample	+/-125	+/- 65	Total sample	+/-125	+/- 65
Loan	0.4683	0.3624	0.4194	0.1731	0.2155	0.1520
	0.1762	*0.1250*	*0.1464*	*0.0751*	*0.0541*	*0.0663*
Empirica	-0.0028	-0.0008	-0.0019	0.0009	-0.0001	0.0009
	0.0024	*0.0009*	*0.0013*	*0.0015*	*0.0005*	*0.0008*
Empirica x above	0.0015	0.0004	0.0013	-0.0002	0.0005	0.0000
	0.0021	*0.0008*	*0.0015*	*0.0014*	*0.0005*	*0.0009*
Empirica sq	-0.0221			0.0078		
	0.0225			*0.0143*		
Empirica sq x above	0.0274			-0.0094		
	0.0246			*0.0149*		
Intercept	0.4358	0.4635	0.4428	0.5446	0.5249	0.5530
	0.0506	*0.0377*	*0.0439*	*0.0360*	*0.0258*	*0.0309*
# observations	2 304	2 007	1 397	6 909	5 817	4 027

	Borrower wage below median (IV)			Borrower wage above median (IV)		
	Total sample	+/-125	+/- 65	Total sample	+/-125	+/- 65
Loan	0.2755	0.2561	0.2725	0.2051	0.2361	0.1604
	0.1212	*0.0815*	*0.1005*	*0.0837*	*0.0626*	*0.0751*
Empirica	-0.0004	0.0000	-0.0004	-0.0001	-0.0006	0.0007
	0.0018	*0.0006*	*0.0010*	*0.0018*	*0.0006*	*0.0010*
Empirica x above	0.0003	0.0000	0.0005	0.0006	0.0009	0.0002
	0.0016	*0.0006*	*0.0010*	*0.0017*	*0.0006*	*0.0011*
Empirica sq	-0.0051			0.0021		
	0.0179			*0.0165*		
Empirica sq x above	0.0055			-0.0027		
	0.0197			*0.0170*		
Intercept	0.4994	0.5053	0.4975	0.5247	0.5108	0.5460
	0.0416	*0.0296*	*0.0352*	*0.0433*	*0.0315*	*0.0376*
# observations	4 607	3 993	2 794	4 606	3 831	2 630

Sample: loans requested for registration in a public University ("Hemis perimeter"). The unit of observation is loan demand per year. Loans requested in November/December are excluded as in all baseline estimations. Explained variable is a dummy that takes value 1 when a the student is found enrolled at a public University on the same year as the loan request. IV estimation: the excluded instrument is a dummy for "above discontinuity point". Robust standard errors in italics.

Table7: University outcomes as a function of loan obtention (Hemis perimeter, 2004-2007)

	Total sample	Total sample	IV for various bandwidth			
	OLS	IV	+/-125	+/-100	+/- 65	+/-20
	Number of courses registered					
Loan	1.0562	1.8599	1.5139	1.7448	1.4295	1.6413
	0.1108	*0.6611*	*0.4800*	*0.5177*	*0.5815*	*1.1290*
Empirica	0.0043	-0.0053	-0.0030	-0.0035	-0.0020	-0.0117
	0.0072	*0.0105*	*0.0039*	*0.0043*	*0.0061*	*0.0266*
Empirica x above	-0.0016	0.0030	0.0075	0.0048	0.0087	0.0148
	0.0091	*0.0098*	*0.0038*	*0.0044*	*0.0068*	*0.0316*
Empirica sq	0.0408	-0.0321				
	0.0791	*0.0973*				
Empirica sq x above	-0.0139	0.0774				
	0.0766	*0.1050*				
Intercept	3.6461	3.3745	3.4163	3.3943	3.4449	3.3191
	0.1155	*0.2471*	*0.1813*	*0.1896*	*0.2148*	*0.3727*
	Credits completed					
Loan	0.0604	0.1189	0.0844	0.0999	0.0716	0.1734
	0.0084	*0.0495*	*0.0363*	*0.0392*	*0.0436*	*0.0851*
Empirica	0.0004	-0.0003	-0.0002	-0.0002	0.0001	-0.0030
	0.0005	*0.0008*	*0.0003*	*0.0003*	*0.0004*	*0.0020*
Empirica x above	-0.0004	0.0000	0.0006	0.0004	0.0003	0.0023
	0.0007	*0.0007*	*0.0003*	*0.0003*	*0.0005*	*0.0025*
Empirica sq	0.0034	-0.0019				
	0.0061	*0.0074*				
Empirica sq x above	0.0006	0.0072				
	0.0059	*0.0080*				
Intercept	0.2353	0.2156	0.2187	0.2168	0.2267	0.1876
	0.0087	*0.0182*	*0.0136*	*0.0141*	*0.0158*	*0.0274*
# observations	9 979	9 979	8 531	7 717	6 011	2 340

Sample: loans requested for registration in a public University ("Hemis perimeter"). The unit of observation is loan demand per year. Loans requested in November/December are excluded as in all baseline estimations. The sample is NOT restricted to individuals enrolled at University. Explained variable are for the same academic year as the loan request. In IV specification, the excluded instrument is a dummy for "above discontinuity point". Robust standard errors in italics. Bandwidth of +/-125 around discontinuity point is the optimal bandwidth according to the cross-validation criteria for both variables.

Table 8: University enrolment as a function of loan obtention, difference-in-difference (Hemis perimeter, 2007-2008)

	Total sample		IV for various bandwidth			
	IV	IV	+/-125	+/- 100	+/- 65	+/-20
Loan	0.1897	0.2048	0.2056	0.2257	0.2735	0.2122
	0.1112	*0.0953*	*0.0978*	*0.1003*	*0.0989*	*0.2122*
Above discontinuity point	*0.0109*	*-0.0128*	*-0.0129*	*-0.0135*	*-0.0273*	*0.0020*
	0.0407	*0.0376*	*0.0368*	*0.0373*	*0.0385*	*0.0750*
Year 2007	*0.0318*	*0.0136*	*0.0087*	*0.0239*	*0.0317*	*0.0536*
	0.0357	*0.0278*	*0.0283*	*0.0289*	*0.0326*	*0.0617*
Empirica x 2007	-0.0003	0.0001	0.0000	0.0001	0.0000	-0.0002
	0.0020	*0.0007*	*0.0007*	*0.0008*	*0.0011*	*0.0048*
Empirica x above x 2007	0.0005	0.0004	0.0005	0.0001	-0.0002	0.0015
	0.0019	*0.0007*	*0.0007*	*0.0008*	*0.0013*	*0.0057*
Empirica sq x 2007	-0.0047					
	0.0195					
Empirica sq x above x 2007	0.0069					
	0.0201					
Empirica x 2008	-0.0006	0.0013	0.0000	0.0001	0.0000	-0.0002
	0.0016	*0.0005*	*0.0007*	*0.0008*	*0.0011*	*0.0048*
Empirica x above x 2008	0.0014	-0.0006	0.0005	0.0001	-0.0002	0.0015
	0.0018	*0.0006*	*0.0007*	*0.0008*	*0.0013*	*0.0057*
Empirica sq x 2008	-0.0202					
	0.0169					
Empirica sq x above x 2008	0.0194					
	0.0168					
Intercept	0.4879	0.5106	0.5130	0.4982	0.4841	0.4549
	0.0297	*0.0209*	*0.0218*	*0.0231*	*0.0266*	*0.0449*
# observations	7 134	7 134	6 169	5 586	4 376	1 746

Sample: loans requested for registration in a public University ("Hemis perimeter"). The unit of observation is loan demand per year. Loans requested in November/December are excluded as in all baseline estimations. Explained variable is a dummy that takes value 1 when a the student is found enrolled at a public University on the same year as the loan request.In IV specification, the excluded instrument is a dummy for "above discontinuity point x year 2007". Robust standard errors in italics.

Table 9: University enrolment as a function of loan obtention

Hemis + Unknown perimeter, 2004-2007

	Total sample		IV for various bandwidth			
	IV	IV	+/-125	+/- 100	+/- 65	+/-20
Above discontinuity point	0.1780	0.1873	0.1926	0.1822	0.1666	0.1791
	0.0728	*0.0431*	*0.0512*	*0.0556*	*0.0640*	*0.1340*
Empirica	0.0004	0.0002	0.0001	0.0002	0.0003	0.0002
	0.0010	*0.0003*	*0.0004*	*0.0004*	*0.0006*	*0.0027*
Empirica x above	0.0000	0.0003	0.0003	0.0004	0.0004	0.0004
	0.0009	*0.0003*	*0.0003*	*0.0004*	*0.0006*	*0.0030*
Empirica sq	0.0026					
	0.0088					
Empirica sq x above	-0.0026					
	0.0097					
Intercept	0.5249	0.5204	0.5183	0.5206	0.5253	0.5221
	0.0246	*0.0168*	*0.0176*	*0.0185*	*0.0213*	*0.0407*
# observations	12 643	12 643	10 945	9 969	7 868	3 117

Sample: loans requested for registration in a public University ("Hemis perimeter") or under unknown status (either public or private University). The unit of observation is loan demand per year. Loans requested in November/December are excluded as in all baseline estimations. In the second panel, explained variable is a dummy that takes value 1 when a loan has been granted. In IV specification, the excluded instrument is a dummy for "above discontinuity point". Robust standard errors in italics.

Table 10: University enrolment as a function of loan obtention (Hemis perimeter, 2004-2007)

Including loan requests in November/December

	Total sample		IV for various bandwidth			
	IV	IV	+/-125	+/-100	+/- 65	+/-20
Loan	0.2294	0.2180	0.2259	0.2232	0.2001	0.2985
	0.0669	*0.0413*	*0.0474*	*0.0514*	*0.0585*	*0.1166*
Empirica	-0.0009	-0.0002	-0.0004	-0.0004	-0.0002	-0.0025
	0.0011	*0.0004*	*0.0004*	*0.0004*	*0.0006*	*0.0029*
Empirica x above	0.0015	0.0005	0.0007	0.0008	0.0009	0.0015
	0.0010	*0.0003*	*0.0004*	*0.0004*	*0.0007*	*0.0033*
Empirica sq	-0.0077					
	0.0102					
Empirica sq x above	0.0064					
	0.0110					
Intercept	0.5047	0.5162	0.5102	0.5109	0.5168	0.4830
	0.0256	*0.0180*	*0.0186*	*0.0194*	*0.0221*	*0.0401*
# observations	11 214	11 214	9 570	8 645	6 715	2 584

Sample: loans requested for registration in a public University ("Hemis perimeter"). The unit of observation is loan demand per year. Explained variable is a dummy that takes value 1 when a the student is found enrolled at a public University on the same year as the loan request. In IV specification, the excluded instrument is a dummy for "above discontinuity point". Robust standard errors in italics.

Table 11: Enrolment at UNISA as a function of loan obtention for individuals who did **not** ask a loan for UNISA
(Hemis perimeter, 2004-2007)

	Total sample		IV for various bandwidth			
	IV	IV	+/-125	+/-100	+/- 65	+/-20
Loan	-0.0003	0.0021	-0.0006	-0.0041	-0.0028	-0.0145
	0.0115	*0.0082*	*0.0089*	*0.0095*	*0.0112*	*0.0190*
Empirica	-0.0002	0.0002	0.0002	0.0001	0.0000	0.0007
	0.0004	*0.0001*	*0.0002*	*0.0002*	*0.0003*	*0.0012*
Empirica x above	-0.0044	-0.0003	-0.0002	0.0000	0.0002	0.0006
	0.0036	*0.0001*	*0.0002*	*0.0002*	*0.0003*	*0.0017*
Empirica sq	0.0004					
	0.0005					
Empirica sq x above	0.0027					
	0.0037					
Intercept	0.0339	0.0397	0.0394	0.0384	0.0356	0.0420
	0.0097	*0.0070*	*0.0071*	*0.0075*	*0.0086*	*0.0144*
# observations	6 893	6 893	5 811	5 222	4 056	1 580

Sample: loans requested for registration in a public University ("Hemis perimeter") other than UNISA. The unit of observation is loan demand per year. Explained variable is a dummy that takes value 1 when a the student is found enrolled at UNISA on the same year as the loan request. The excluded instrument is a dummy for "above discontinuity point". Robust standard errors in italics.