Should Program Graduation be Better Targeted? The Other Schooling Outcomes of Mexico's Oportunidades*

Tobias Pfutze†
Florida International University

Abstract

A large literature on Conditional Cash Transfers programs assesses the effects of becoming a beneficiary. However, the consequences of losing the benefit due to program graduation are largely unstudied. This paper replicates the eligibility score employed over 2010-15 by Mexico's Oportunidades for a very large household survey. Using a Regression Discontinuity Design around the threshold for program graduation, it shows that losing this additional incentive had a negative effect on high school attendance for lower secondary school aged students in urban, and upper secondary school aged ones in rural areas. The results suggest that the graduation thresholds are possibly too low given the program's stated aims.

JEL Classification: I25, I38, J22, O12
Keywords: Education, Conditional Cash Transfer Program, Gender, Mexico

*The findings, interpretations, and conclusions expressed in this work are exclusively the author's responsibility, and do not form part of the official statistics or positions of Mexico's National Institute of Statistics and Geography (INEGI), nor of the National System of Statistical and Geographical Information (SNIEG)

†Department of Economics; Florida International University; 11200 SW 8th Street; Deuxième Maison(DM) 315B; Miami, FL 33199; Phone: (305) 358-2381; Email: tpfutze@fiu.edu

1 Introduction

Since the inception of Mexico's Progresa in 1997, cash transfer programs, either conditional or unconditional, have been one of the primary areas of research in the development literature. The focus has been almost exclusively on the various effects of program implementation, or becoming a beneficiary, on a variety of outcomes on the individual, household, or some aggregate unit (e.g. the village). While this is undoubtedly the single most important treatment, it largely ignores that beneficiaries may drop out of the program as their socioeconomic circumstances, possibly in part as a result of the program itself, improve. Governments have to strike a delicate balance between, on the one hand, control expenditures by graduating beneficiaries who no longer qualify from social programs; and, on the other, make sure that the loss of benefit does not eradicate the gains made.

This paper is to my knowledge the first study that looks at the effects of program graduation on one of the target outcomes of conditional cash transfer programs, namely school attendance. It does so for the well-studied Mexican flagship program Oportunidades. The findings show that program graduation significantly reduces school attendance at the lower secondary school level (grades 7-9) for urban students and at the upper secondary level (grades 10-12) for rural ones. Moreover, there is evidence, albeit weaker, that male students after dropping out do more work in the household, while female ones are more likely to neither work nor study.

The paper employs a regression discontinuity design (RD) around the known threshold for program graduation, based on individual and household level data collected by Mexico's 2015 Intercensal Survey (*Encuesta Intercensal*). This is combined with information on the recertification process during 2011/12. Since the data does not contain information on a household's beneficiary status, the results presented have to be interpreted as an intention to treat effect, rather than an average treatment effect. By this measure, lower secondary school attendance is reduced by around five percentage points for urban students, and over 14 percentage points for rural students at the upper secondary level. However, when combined with other information on program coverage, the estimated average treatment effects increase to 26% and 16%, respectively. Given the general characteristics of the program's implementation in the late 1990s and early 2000s, the vast majority of households in the sample can be expected to have been beneficiaries since the first half of the 2000s or even since the 1990s. The results, therefore, strongly suggest that the conditional part of the cash-transfer program continues to play an important role even for households who have benefitted from it for a long period of time. In this context, the study also makes an important contribution to the emerging literature on the longer term effects of cash-transfer programs.

In order to keep the language simple, I will refer to the program simply as Oportunidades. As the reader may be well aware, it was initially instituted under the name Progresa, which it held until after the change in Mexico's federal government in late 2000. In 2014, over a year after the old governing party PRI returned to power, the name was changed yet again to Prospera. I opt for the name Oportunidades, since the treatment of interest (benefit loss due to recertification) occurred during 2011 and 2012 when the program was still operating under that name.

In the next section, I will provide a description of the program itself and the recertification process. This is followed, in the same section, by a review of the literature relevant to this study. Section three discussed the RD methodology employed, and section four explains the data used and how the methodology is implemented in the present setting. Section five then discusses the results and section six concludes.

2 Background and Literature

Mexico's Conditional Cash Transfer program Oportunidades first began operating in 1997 (under the name Progresa) in a few rural localities. In 1998, 506 additional localities were selected, of which 320 were randomly assigned to treatment, while the remaining 186 localities entered the program 18 months later. It was this random protocol that made it the probably most researched and well documented government program outside high-income countries. Levy (2006) gives an in-depths about the creation of the program and an exhaustive review of its earlier literature. In a nutshell, Oportunidades provides cash payments to families conditional on children's school attendance and that all family members visit health clinics on a regular basis. First solely focused on rural areas and children of mandatory school age, in 2001 the program started expanding into urban areas and to provide additional subsidies for children attending school beyond ninth grade. While the program's name changed with the two changes of party in government, it maintained its principal characteristics with relatively minor changes. In order to keep the discussion to a manageable length, at continuation I will only go over the program rules and the more recent literature directly relevant to this study.

At the core of the analysis that follows is the process followed to determine graduation from the program. For Oportunidades, eligibility is determined through an estimated per-capita household income based on a number of easily observed household characteristics captured by a questionnaire called *ENCASEH* (Encuesta de Características Socioeconómicas y Demográficas de los Hogares) administered during visits to the household. If the estimated income is below a pre-defined minimum welfare line (LBM, by its Spanish acronym) the household qualifies. For recertification, households' eligibility was supposed to be reassessed at first every three years. If the household continued to be below the LBM, the program was continued. If it was above the LBM, but still below a second, higher, line, the household was put into a regime with reduced benefits. Importantly, for what follows, the stipends for post-primary education were not affected. If the household moved above the second line, benefits were terminated. The basic setup of this process has not changed since. However, the three yearly recertification proved to be excessively burdensome since it required that each year one-third of beneficiaries, spread out over the entire country, be interviewed. In 2010, in addition to updating the underlying model used to determine eligibility and permanence to the one used in this study[1], the recertification process was changed to a five yearly interval. Now, each year a determined number of localities (instead of households), comprising around one-fifth of the beneficiary population, are visited. The order of visits was determined by localities' score on the Social Gap Index (*Índice de Rezago Social*), calculated by Mexico's National Council for the Evaluation of Social Policies (*CONEVAL*). This index gives each locality in the country a social gap score (very low, low, medium, high, very high) according to its performance on a multi-dimensional poverty measure. Households in localities with the lowest gap were put through recertification first.

This process, however, proved to be highly controversial as during the first two years around 30% of households that went through it were graduated from the program. It was stopped in late 2012 when a new federal administration entered office. Even though high, the graduation rate of 30% does not seem unreasonable. In my conversations with government officials working for the program, it was mostly attributed to comparatively high level of social development in these localities. The model used to deter-

[1]The model used during late 2010-15 was based on information collected in the 2010 *Encuesta Nacional de Ingresos y Gastos de los Hogares* (ENIGH).

mine eligibility was again changed in 2015 (based on the 2014 round of the ENIGH), and recertification was reinstated at the locality level, but without the grouping by the Social Gap Index. Currently, around 16% of households are dropped from the program each year. The upshot for the present paper is that the results presented apply for localities with relatively high socioeconomic development (84% of observations in the mostly used sample live in municipalities[2] with very low or low gaps).

The model used to estimate household per-capita income during 2010-15 consists of a linear regression of the logarithm of per-capita household income (excluding government transfers). Different models, based slightly different characteristics are estimated for rural and urban areas; where rural, following the standard definition used in Mexico, refers to localities with less that 2,500 inhabitants. The characteristics of a household's dwelling and its integrants are the following: Food security (with different measures for rural and urban areas), having a toilet with water connection, solid floors (only rural), covered floors, home ownership (only urban), paying rent, number of rooms, number of women aged 15-49 (only urban), logarithm of number of household members, proportion of dependents, schooling of household head and spouse, number of household members working as subordinates, number of household members working as independents, number of household members working without pay (only urban), not having a refrigerator, not having a motor vehicle, not having a computer (only urban), not having a

[2]The locality level index is not observed

VHS or DVD player, not having a landline phone (only urban), not having a microwave oven, using wood/coal/oil as cooking fuel, being a Seguro Popular beneficiary (Mexico's non-contributory health insurance, only urban), at least one household member having employer provided health insurance (with and without household head working independently), municipal index of social underdevelopment, household receives remittances, and locality size 15,000-99,999/more than 100,000 inhabitants[3].

The literature on Oportunidades is plentiful. Even the original randomized setup continues to spawn new studies after more than 15 years. Recent examples include Alix-Garcia, McIntosh, Sims & Welch (2013), who show that the additional income accelerated deforestation, particularly in areas with poor infrastructure and market access. Behrman & Parker (2013) find that the program also had a positive health effect on the elderly living in beneficiary households, especially for women. Recent work by Gertler, Martinez & Rubio-Codinsa (2012) show that households use part of their benefit to invest in productive asset, increasing its longer-term impact on consumption. Dubois & Rubio-Codina (2012) argue that the program also increases human capital by allowing mothers to spend more time with their young children. Not using the original randomization, which was restricted to rural localities, but nonetheless of interest, Behrman, Gallardo-Garccía, Parker, Tood & Vélez-Grajales (2012) show that the program's expansion to urban areas

[3]The appendix to this paper shows the full probit model with its corresponding parameter values

had similar effects to those found in rural settings. Specifically, it increased school enrollment and attainment, increased time devoted to homework for girls and working rates of boys. The interest in CCTs is not limited to Mexico. Important recent studies on other countries can be found for Malawi (Baird, McIntosh & Özler 2016), or Brazil's *Bolsa Escola/Familia* (Glewwe & Kassouf 2012). The former case is of interest, since it only ran for two years in 2008/09. Using data through 2012, the authors find that after the program has been discontinued, its positive impact quickly dissipated. The second paper uses data from Brazil's School Census over the 1998-2005 period, to show that the program increased enrollment and grade promotion, while lowering dropout rates.

One of the first studies to assess the longer term effects of Oportunidades is Behrman, Parker & Todd (2011), considering a time frame of roughly five and a half years. Given the focus on schooling and work, it is closely related to the present study. The authors first compare outcomes in 2003 in response to the 18 month lag in exposure due to the randomized roll-out. The authors find significant increases in school performance for boys and girls aged 9-15 in 1997, equivalent to about one-fifth of a grade. The results for work are somewhat weaker, but still indicate some reduction for boys in the same age group. The second exercise consists of comparing each of these two groups (who had either five and a half years or four years of program exposure in 2003) to a control group that had never been exposed, using difference in differences matching estimators. The results show a strongly significant increase in school performance for both groups, but a larger effect (roughly 0.7 to a full grade vs. 0.5 to 0.75) for the group with longer exposure. The only significant results for work and agricultural work were found for boys aged 15-16, for who that probability was reduced by 14% and 9%, respectively. A similar exercise by the same authors (Behrman, Parker & Todd 2009) for younger children who were not of school age at the start of the program, but should have benefitted from the nutritional and health components, shows that it significantly reduced the age of primary school entry. Looking at a somewhat longer term, Rodríguez-Oreggia & Freije (2016) look at similar outcomes in 2007 for beneficiaries who were 5-15 years old in 1998. They do not find results on employment, wages or inter-generational mobility. That said, the sample employed suffers from large rates of attrition. The three previous paper used Mexico's household evaluation survey (ENCEL). Parker, Rubalcava & Teruel (2012) instead employ the MxFLS, which, unlike the ENCEL, is representative for the whole country. Using the survey's 2002, 2005, and 2009 rounds, the sample consists of individuals who were 10-14 years old in 1997. Results are presented for difference in differences estimates with a variety of different matching methods, showing a consistently big (around 5 percentage points) and significant increase of the probability of attending college. The labor market results are more mixed, with some evidence of a higher probability of working, but no results regarding hours worked, or wages and benefits received. Looking further into the future, McKee & Todd (2011) use the existing evidence of the program's human capital effects, and data on current 25-40 year olds, to simulate its long-term effects on earnings. They conclude that while its does increase mean earnings, it will only have limited effects on earnings inequality.

Moving beyond Mexico, long-term effects of other CCT programs in Latin America have also been assessed for Nicaragua's *Red de Protección Social*, Colombia's *Familias en Acción*, and Ecuador's *Bono Desarrollo Humano.* The Nicaraguan case is of interest because localities were randomly assigned to either receiving the benefit over a three year period starting in either 2000 or 2003. It therefore allows for the comparison of two distinct groups who had received the benefit at different points in time, but ceased to do so. Barham, Macours & Maluccio (2016) do this for boys aged 9-12 in 2000 with data collected over the 2009-11 period, i.e. after individuals in the sample can be expected to have finished school. Given program setup, the control group benefitted from the program only when aged 12-15, which, as argued by the authors, is too late to make an important difference. It is found that early exposure increased total schooling by half a grade, and also very significantly increased performance in standardized test scores. For the Colombian program, Baez & Camacho (2011) conduct two different exercises, with different datasets, for children who could be expected to finish grade 11 by 2009. The first one consists of a difference in differences estimation based on matching, and the second, somewhat mirroring the strategy followed here, on a regression discontinuity design based on the eligibility score (*Sisben*). The results from both methods show a significant positive impact on school completion, but no effect on test scores. While the RD design shows a consistent increase on the first count for all subgroups (with increases between 2 and 8 percentage points), the matching estimators only show a positive effect for females and in the rural sector (with point estimates sometimes above ten 10 percentage points). A similar approach is taken by Araujo, Bosch & Schady (2016), who show that Ecuador's program, after ten years, had no effect on test scores, but increased secondary school completion for females by around two percentage points. For the first result, the authors employ randomization during early implementation which determined the number of years a child has benefitted from the program. The second set of results is obtained by regression discontinuity analysis.

So while this is not the first paper to look at a CCTs longer-term effects, there is a surprising lack of literature on program graduation. This is true for the determinants of graduation, as well as, its consequences. The to my knowledge only paper on this topic is Villa & Nino-Zarazúa (2014), who use the Mexican Family Life Survey (MxFLS), a longitudinal dataset with a first round collected in 2002 and follow-up rounds in 2005/06 and 2011/12, to estimate graduation probabilities. The models are run on a large number of observable household characteristics, separately for rural and urban areas. The authors conclude that the program was able to graduate 28.9% of its urban, and 26.7% of its rural beneficiaries. Focusing not on graduation, but on dropout, Álvarez, Devoto & Winters (2008) argue that the program's conditionality imposes higher compliance costs on relatively better-off households, increasing their likelihood to drop out. This increased overall program targeting on the poor.

3 Methodology

First proposed by Thistlewaite & Campbell (1960), regression discontinuity designs have gained prominence in economics starting in the 1990s. Over the years, a large variety of different RD estimators have been proposed. The approach used here consists of using nonparametric regressions to determine the expected values of the outcome of interest at the boundary points on either side of the threshold, conditional on the assignment variable. The treatment effect is then calculated as the difference between the two. This method has gained prominence since Hahn, Todd & Van-der Klaauw (2001), given the problems with standard nonparametric kernel regressions in this context, proposed local regressions[4] to estimate the boundary points. As of the writing of this paper, this has arguably become the most common approach.

While Hahn et al. (2001) showed that the local regression is less biased than other nonparametric methods, it only yields consistent estimates under fairly strong assumptions. Of crucial importance here is the bandwidth employed in the estimator: While a larger bandwidth will result in more

[4] Local regressions were first proposed by Fan (1992).

precise estimates (since it is based on more observations), it also increases the potential bias. This tradeoff between bias and precision is central to all nonparametric estimators. This study employs a nonparametric local linear regression, including the recent advances proposed by Calonico, Cattaneo & Titinuk (2014*b*). In order to establish robustness of the results, it follows the recommendations in Lee & Lemieux (2010) showing results for different neighborhoods, and polynomials. It will also show the distribution of the assignment variable, give a graphical impression of the discontinuity, and show the continuity of other potential observable determinants of the outcome.

To fix ideas, following Calonico et al. (2014*b*) the sharp RD estimator can be expressed as:

$$b^S_{RD} = \alpha^+_y - \alpha^-_y, \tag{1}$$

where the y and subscript denotes the outcome of interest, and the superscripts refer to the boundary values at the threshold from the left (-) and right(+). The value of α^+_y is obtained as the estimator $\hat{a}_y$ of the local linear regressions:

$$(\hat{a}_y, \hat{c}_y) = \underset{a_y, c_y}{argmin} \sum_{i=1}^{N} \{y_i - a_y - c_y(x_i - \tau)\}^2 K\left(\frac{x_i - \tau}{h}\right) I[x_i \geq \tau]$$

Here, τ denotes the threshold, and $K(.)$ a kernel function of a type to be chosen by the researcher. N are all observations in the data or a neighborhood around the threshold, and h is the bandwidth (which, of course, allows for some observations to be completely excluded). α_y^- is obtained similarly to the left of the threshold, i.e. changing the identity function to $I[x_i < \tau]$.

Calonico, Cattaneo & Titinuk (2014*a*) provide a detailed overview of the different methods available and how the relate to the refinements proposed by the authors; the aforementioned Calonico et al. (2014*b*) provides a technical discussion of their methods. Traditionally, the biases inherent to the point estimator and its confidence interval were either ignored, which may be justified when employing sufficiently small bandwidths, or by estimating a bias correction based on a local regression with a polynomial at least one order higher than the one used in the primary local regression. Calonico et al. (2014*a*) conduct a number of Monte Carlo experiments to demonstrate several shortcomings in the estimators discussed thus far, and to show the better performance of their proposed method in finite samples. In a nutshell, their principal contribution is that instead of subtracting the estimated bias term from the confidence intervals, they derive the asymptotic variance for the bias-corrected point estimate. This variance term takes into account the additional variation introduced by the correction, allowing for robust standard errors and thus confidence intervals with improved finite sample properties. The results presented here will show the bias corrected point estimates and the t-statistic based on robust standard errors.

4 Data and Implementation

The principal data source employed is Mexico's 2015 Intercensal Survey (INEGI 2015), collected between March 2-27, 2015. Since the 19th century, Mexico conducted a full census in all years ending in zero, with the occasional deviation from this pattern due to internal or external strife. In 1995 and 2005, the country also carried out and additional census (called *Conteo*) with a slightly shorter questionnaire. This practice was again changed in 2015, when instead of a full census the National Institute of Statistics and Geography (*INEGI*, by its Spanish acronym) decided conduct a large household survey with a detailed questionnaire[5]. The Encuesta Intercensal has a sample size of 6.1 million households, designed to be representative of all localities with more than 50,000 inhabitants.

The single most important reason to use the Encuesta is its large sample size. Results will be presented for different groups, in particular rural males and females, as well as, their urban counterparts. The large sample size provides sufficient thickness of observations around the threshold in each group to draw meaningful conclusions. Empirical researchers often face an important trade-off between large sample sizes and more observable characteristics, as smaller surveys can afford to administer larger and more detailed questionnaires. The present study is no exception as the large sample comes

[5]The questionnaire follows the same design as the extended questionnaires used in earlier censuses for a subsample of the population.

at the price of not being able to observe a household's beneficiary status. The largest survey available that would contain all the relevant information is the already mentioned ENIGH, which is collected in all even years. However, its comparatively small sample size (the 2014 round contains a total of 21,400 households) renders RD estimation infeasible. The estimations are, therefore, conducted as a sharp RD based on the permanency threshold. This approach can be interpreted as an intention to treat effect, rather than an average treatment effect. For a fuzzy RD estimator the expression in (1) would be divided by the (similarly estimated) change in the treatment at the threshold:

$$b_{RD}^{F} = \frac{\alpha_y^+ - \alpha_y^-}{\alpha_d^+ - \alpha_d^-}, \tag{2}$$

The denominator of which has be be between zero and one. The estimates presented can thus be thought of as a lower bound on the true treatment effect. In this sense, the lack on information on beneficiary status works against the econometrician by biasing the estimates towards zero.

The group of interest for this study are individuals aged 13-15 and 16-18, respectively. The typical Mexican student enters lower secondary school (*Secundaria*) after six years of primary school at age 12, and upper secondary school (*Preparatoria*) at age 15 and graduates at age 18. The data contains the age of each person surveyed as of March 2015, but not the exact date of birth. But since they were collected about 3 months before the end of the school year the majority of students in 12th grade can be assumed to be 18 years old at that point. Likewise, most students in 10th grade can be assumed to be already 16 years of age; most of those in 9th, 15 years; and those in 7th, 13 years. The sample is restricted to children living in localities that participated in the recertification process during 2011/12[6]. The sample is further restricted to those children who have met the minimum requirements to attend lower and upper secondary school. That is, those that have successfully graduated from primary school and lower secondary school, respectively. It also excludes children who are not living in the household of their parents.

There are two additional sources of noise that will increase the "fuzziness" around the threshold, and hence work against finding any significant effects. The first one is the time lag between recertification in 2011/12 and observation in 2015. However, the household characteristics used can be expected not to change much over 3-4 years (see the discussion above). Households' demographic characteristics have been adjusted to take this time lag into account.[7] On the other hand, some time lag between recertification and observation seems necessary to obtain any significant results, since parents are unlikely to pull their children out of school the day they lose the benefit. Secondly, the Encuesta has information, with almost identical wording as in the

[6] I use ENCASEH data on the universe of recertified household to determine which localities took part

[7] For most localities, the data shows that almost all households went through recertification in the same year. For the localities that had interviews conducted in both years, the year with most interviews was taken used to adjust the demographic characteristics.

ENCASEH, on all these item with the exception of whether the household has a VHS or DVD player. In order to make up for this lack of information, predicted values for DVD/VHS ownership are constructed from characteristics observable in the ENIGH and the Encuesta Intercensal, using the former to estimate a probit model applied to the latter. The explanatory variables used are dwelling characteristics and ownership of durable goods and other services (with the exception of characteristics employed in the index itself)[8]. This correction, while not ideal, is not expected to affect the results in any significant manner. The weight given to DVD/VHS ownership in the index is fairly small, and will therefore only add some additional noise to the index.

The index is is designed as a proxy measure for per-capita household income, based on easily observable household characteristics. The parameters were estimated by regressing the logarithm of income, as observed in the ENIGH, on the characteristics discussed above. The index itself is defined over the predicted monthly per-capita income in 2010 Mexican Pesos (MXN) (i.e. by using the predicted value from the linear model in an exponential function). In order to qualify for the benefit, an urban household had to score less than 1,243.15 MXN, and a rural one less than 716.17 MXN. This paper, however, does not use the qualification threshold for new beneficiaries, but the second, higher, threshold used to determine whether the benefit is discontinued. The corresponding values here are 1,538.29 MXN for urban, and 1,145.65 MXN for rural households. Using the index and these latter

[8] Regression results of the model are presented in the appendix.

values, each household's predicted distance to the cutoff is computed. It is this distance measure that will be at the heart of the analysis that follows. Figure A.2 show as histogram of this distance for all minors aged 16-18 in the sample, up to a distance of 1,000 MXN. Given how the index is constructed, and that it is meant to proxy for income, it is not surprising that it follows a log-normal distribution (it would be worrisome if it did not). The graph shows about 80% of all observations, with the remainder forming the cut-off tail[9]. There is no visible jump or any other discontinuity at the threshold. Given the distribution, there will always be some more observations to its left than to its right.

The principal specification used in the analysis below uses observations within a +/-100 MXN neighborhood around the permanency threshold, combined with a local linear regression. However, in light of the discussion above, results will also be presented for wider neighborhood and higher local polynomial regressions. Taking a closer look at the the different outcomes of interest, table 1 shows summary statistics for the eight different groups of minors aged 13-15 and 16-18 within +/-100 MXN of the permanency threshold. The first line in each of the four panels shows the outcome of school attendance at their corresponding level. On average, around 90% of those aged 13-15 attend lower secondary school, and 69% of males and 75% of females aged 16-18 upper secondary. The next four outcomes capture a per-

[9]The survey stratification over samples households in smaller, and on average poorer, localities

son's primary occupation during the week prior to the interview. There are always a few percent fewer children who primarily study than are attending school. This indicates that some of those who do attend school, do not consider this their primary activity. Unsurprisingly, more of the older children are primarily in paid work. Also, more males work than females, while the latter are more commonly found in household chores. Close to 10% of older, and 5%-6% of younger, males do neither work nor study. This number is slightly lower for females. The remaining characteristics will be used to determine smoothness. They are capture whether or not a child is fully or partially indigenous; his/her parents average years of education (measured by the household and and spouse, and only the former if no spouse is present); a binary variable indicating that no spouse or domestic partner of the household head is present; the number of other minors in the household; a binary variable indicating a nuclear family household; the size of the household; the age of the household head; a binary variable indicating that at least one member owns agricultural or grazing land; a binary variable indicating that the household suffers from food vulnerability[10]; and reported household income excluding transfers. The number should be self-explanatory, and contain no real surprises.

[10]defined someone in the household having to forgo meals or eat less for lack of economic resources

5 Results

I start by presenting results for the local linear regression RD estimates on school attendance within a +/- 100 MXN neighborhood around the eligibility threshold. For each age group, results will be presented for the full sample, urban vs. rural, male vs. female, and lastly all four sub groups separately. This is followed by an analysis of the impact on primary activities for the groups with significant results on school attendance. After that, I will present and discuss various robustness tests. While results for school attendance differ between rural and urban, they are fairly similar for males and female. The robustness part, therefore, pools the two genders. I will present results for smoothness of other variables around the threshold. This is followed by estimations using different data windows around the threshold and local linear polynomials of first, second, and third order. Finally, I will give a visual impression of the results.

The tables for RD results, based on local regressions, show the estimated size of the bias-corrected discontinuity at the threshold, its robust p-value, the number of observations effectively used at the left and right of the threshold, plus the bandwidths used in the regression and bias correction, respectively. Unless otherwise noted, linear local regressions were conducted. The order of the local polynomial used for the bias correction is always one order higher than that used in the regression.

5.1 Main results

Table 2 shows the most important results: Loss of Oportunidades leads to a statistically significant drop (at the one-percent level with robust standard errors) in school attendance for urban 13-15 year olds and rural 16-18 year olds. The effect on rural areas is of much larger magnitude, and still significant at the one-percent level for females only. This is all the more important given that rural sample sizes are much smaller than urban ones. Probably for the same reason, the urban effect for 13-15 year olds also yields statistically significant results for the whole sample. The point estimates for males and females are very similar. However, the results have lower p-values for male urbanites in the younger cohort compared to the corresponding females (this also yields statistical significance for the pooled sample of males, but not for females). This relationship is inverted for the older cohort in rural areas.

The interpretation of the actual magnitude of the effects is a bit more complicated. The point estimates indicate a drop of five percentage points for the comparatively younger cohort in urban areas and of over 14 percentage points for the older rural students. In order to get an idea of the size of the actual treatment effect, one would need estimates of the magnitude of the denominator in expression 2 above. If it can be assumed that all (or close to all) households to the right of the threshold living in localities that went through recertification were effectively dropped from the program, the overall proportion of beneficiary households to the left can be used as a (rough) estimate of the denominator. These proportions can be calculated with high degree of accuracy by taking advantage of the large sample size in the Encuesta Intercensal 2015, and combining it with official program data. According to the latter, 67,304 household which scored within a distance of MXN 100 to the permanency threshold from below went through recertification during 2011/12 in urban localities, and 52,139 did in rural ones. Using information from the Encuesta, and applying sampling weights, the total number of households within this distance, i.e. including those who were not beneficiaries, should have been 313,079 and 61,209, respectively. This implies pre-certification levels of coverage in this bracket of 21.5% and 85.18%. This would yield a treatment effect in urban areas of 26.37% and 16.28% in rural ones. While these number need, of course, to be taken with a big grain of salt, they do contain some useful information: Firstly, the size of the effect on younger urban, and older rural students is likely much more similar than the point estimates suggest, with a larger actual treatment effect in urban areas compared to rural ones.. Secondly, the magnitude of the effect is indeed important.

Table 3 shows results for a child's main activity in the week prior to the survey interview. Results are only shown for the groups that had significant reductions in school attendance. The outcomes here are necessarily much noisier than the more clear-cut outcome of school attendance, mainly because the precise nature of one's principal activity may be somewhat ambiguous. For that reason, estimations results are also less significant. This problem appears to be particularly severe for paid work. However, some insights can be gleaned. Firstly, for urban males aged 13-15 and rural females, the point estimates found on studying are very close to those for school attendance, albeit at a somewhat lower statistical significance level. It is more difficult to say to which primary activities the drop-outs shifted. For younger urban males, the small shift towards household work (in which generally few males participate) is almost significant at the five-percent level. However, the point estimates for paid work or not working (nor studying) are larger in magnitude, though not significant. Something similar is happening to older rural males, for which the only statistically significant result (at the ten-percent level) is also household work, yet by far the largest point estimate is for shifts to paid work. For rural females the only significant shift is towards neither working nor studying (at the five-percent level), though paid work and household work have not much smaller (yet insignificant) point estimates. For the younger, urban females, the biggest shift is to paid work with a p-value of 0.12.

5.2 Robustness

With the main results firmly established, I will now directly address the assumptions underlying the RD design. I will focus on the results for school attendance and, in the interest of space, male and female students will be pooled, given the similarity in their results. Lastly I will only look at the two groups for which significant results on school attendance were found. Table 4 runs a local linear regression on a number of other observable characteristics to assess smoothness at the threshold (see discussion for table 1 above). For the older, rural students,there is no significant jump in any of these, and p-values are always far removed from anything resembling statistical significance. In urban areas, the p-values are somewhat lower, which is likely attributable to the larger sample size. For two outcomes, the number of other minors and the age of the household head, they are significant at the ten-percent level. None of this cast any doubts on the results found on schooling, which were significant at the one-percent level. Moreover, by showing results for 20 different estimations, finding two of them significant at the ten-percent level is exactly what would be expected.

In table 5, the analysis is extended to different data windows and higher order local polynomials. The size of the data window is halved, to +/-MXN50, and doubled to +/-MXN200. Polynomials are estimated up to the third order. The estimates in the uppermost panel for the +/-MXN100 window are of course the same as presented in table 2. The results are fairly robust in terms of the estimated magnitude of the effect at the threshold. Using a higher order local polynomial on a narrower data window compromises statistical significance, but does not alter the point estimates by much. Wider data windows, in turn, may require a higher order polynomial to yield more precise results (this is the case for urban areas, where the extension of the window results in a much larger bandwidth compared to the rural case).

Lastly, figures 2 and 2 provide a visual impression of the discontinuity, based on a third order polynomial. The much larger point estimate for rural students is clearly visible. There is nothing in the shape of the estimated function that would raise any concerns.

6 Discussion and Conclusions

The analysis just presented showed that graduation from Mexico's flagship social protection program Oportunidades during the 2011/12 recertification cycle resulted in a significant reduction in school attendance to lower secondary school in urban, and to upper secondary in rural areas. While the analysis does not allow for a precise estimation of the treatment effect, some back of the envelope adjustments suggest that in urban areas around 26% of affected 13-15 year olds drop out of school if the benefit is discontinued. The corresponding estimates for rural 16-18 year olds are over 16%.

These results, first and foremost, suggest that given the program's stated aims to foster school attendance in low-income households, the graduation thresholds have been chosen too low. Many households need the incentive of the Oportunidades payment to keep their children in school (or to encourage them to attend). The open question here is, of course, if for the children who drop out in response to the benefit cuts further school attendance would have translated into effective learning outcomes and better job market opportunities in the future. While this question is well beyond the scope of this paper, the observed overall drop still has policy relevance.

Why are there no significant results for younger rural and/or older urban students? The results for these groups in table 2 have very low point estimates and large p-values, suggesting that there is indeed no effect. The most obvious answer here is opportunity costs.Wages that can be earned by a minor are likely to be increasing in his/her age and to be higher in urban areas than in rural ones. If parents do not value their childrens' education prospects, once their earnings potential crosses some critical value, they will opt to take them out of school. Such households would never put their children in upper secondary school independent of the potential Oportunidades benefit, if a 16 year old in an urban setting can earn more than the value of the benefit. Losing the benefit would thus have no additional effect. However, losing the benefit while the child is still in lower secondary school, and has hence a lower earnings potential, would result in an additional drop out. Likewise, for rural households, a lack of job market opportunities for younger students would imply very low opportunity costs of lower secondary school attendance. However, for 16-18 year olds the size of the benefit may be enough to dissuade parents from taking their children out of school, but they would do so in its absence. As has been discussed, the results presented here are based on the comparatively better-off localities. Can they be expected to have external validity and translate to the country at large? If opportunity costs are the primary driver behind the results, the effects should be lower in localities with a larger social gap that offer fewer opportunities.

Future research should further hone in on these questions. Since the effects of program graduation have been largely neglected by the literature there is also scope for many other outcomes of interest, including consumption patterns, poverty risks, or geographical mobility.

References

Alix-Garcia, J., McIntosh, C., Sims, C. R. E. & Welch, J. R. (2013), 'The ecological footprint of poverty alleviation: Evidence from mexico's oportunidades program.', *The Review of Economics and Statistics* **95**(2), 417–435.

Álvarez, C., Devoto, F. & Winters, P. (2008), 'Why do beneficiaries leave the safety net in mexico? a study of the effects of conditionality on dropouts', *World Development* **36**(4), 641?658.

Araujo, M. C., Bosch, M. & Schady, N. (2016), Can cash transfers help households escape an inter-generational poverty trap? NBER Working Paper 22670.

Baez, J. E. & Camacho, A. (2011), Assessing the long-term effects of conditional cash transfers on human capital: Evidence from colombia. IZA Discussion Paper No.5751, Institute for the Study of Labor, Bonn, Germany.

Baird, S., McIntosh, C. & Özler, B. (2016), When the money runs out: Do cash transfers have sustained effects? Unpublished Working Paper.

Barham, T., Macours, K. & Maluccio, J. A. (2016), More schooling and more learning? effects of a three-year conditional cash transfer program in nicaragua after 10 years, Technical report, Inter-American Development Bank. IDB Working Paper No. 432.

Behrman, J. R., Gallardo-Garccía, J., Parker, S. W., Tood, P. E. & Vélez-Grajales, V. (2012), 'Are conditional cash transfers effective in urban areas? evidence from mexico', *Education Economics* **20**(3), 233?259.

Behrman, J. R. & Parker, S. W. (2013), 'Is health of the aging improved by conditional cash transfer programs? evidence from mexico', *Demography* **50**, 1363?1386.

Behrman, J. R., Parker, S. W. & Todd, P. E. (2009), 'Schooling impacts of conditional cash transfers on young children: Evidence from mexico', *Economic Development and Cultural Change* **57**(3), 439–477.

Behrman, J. R., Parker, S. W. & Todd, P. E. (2011), 'Do conditional cash transfers for schooling generate lasting benefits? a five-year followup of progresa/oportunidades.', *The Journal of Human Resources* **46**(1), 93–122.

Calonico, S., Cattaneo, M. D. & Titinuk, R. (2014*a*), 'Robust data-driven inference the regression-discontinuity design', *The Stata Journal* **14**(4), 909–946.

Calonico, S., Cattaneo, M. D. & Titinuk, R. (2014*b*), 'Robust nonparametric confidence intervals for regression discontinuity designs', *Econometrica* **82**(6), 2295?2326.

Dubois, P. & Rubio-Codina, M. (2012), 'Child care provision: Semiparametric evidence from a randomized experiment in mexico', *Annals of Economics and Statistics* **105/106**, 155–184.

Fan, J. (1992), 'Design-adaptive nonparametric regression', *Journal of the American Statistical Association* **87**(420), 998–1004.

Gertler, P. J., Martinez, S. W. & Rubio-Codinsa, M. (2012), 'Investing cash transfers to raise long-term living standards', *American Economic Journal: Applied Economics* **4**(1), 164?192.

Glewwe, P. & Kassouf, A. L. (2012), 'The impact of the bolsa escola/familia conditional cash transfer program on enrollment, dropout rates and grade promotion in brazil', *Journal of Development Economics* **97**, 505–517.

Hahn, J., Todd, P. & Van-der Klaauw, W. (2001), 'Identification and estimation of treatment effects with a regression-discontinuity design', *Econometrica* **69**(1), 201–209.

INEGI (2015), 'Encuesta intercensal 2015'. Access to and use of the microdata was facilitated by INEGI's Microdata Laboratory.

Lee, D. S. & Lemieux, T. (2010), 'Regression discontinuity designs in economics', *Journal of Economic Literature* **48**, 281?355.

Levy, S. (2006), *Progress Against Poverty: Sustaining Mexico's Progresa-Oportunidades Program*, Brookings Institution Press, Washington, D.C.

McKee, D. & Todd, P. E. (2011), 'The longer-term effects of human capital enrichment programs on poverty and inequality: Oportunidades in mexico', *Estudios de Economia* **38**(1), 67?100.

Parker, S. W., Rubalcava, L. N. & Teruel, G. M. (2012), Do conditional cash transfer programs improve work and earnings among its youth beneficiaries? evidence after a decade of a mexican cash transfer program. Unpublished Working Paper.

Rodríguez-Oreggia, E. & Freije, S. (2016), Long term impact of a cash-transfers program on labor outcomes of the rural youth in mexico. Center for International Development at Harvard University Working Paper No. 230.

Thistlewaite, D. L. & Campbell, D. T. (1960), 'Regression-discontinuity analysis: an alternative to the ex-post facto experiment', *Journal of Educational Psychology* **51**, 309–317.

Villa, J. M. & Nino-Zarazúa (2014), Poverty dynamics and programme graduation from social protection: A transitional model for mexico's oportunidades programme. WIDER Working Paper 2014/109.

Figure 1: Distribution of the distance to the permanency threshold by household.

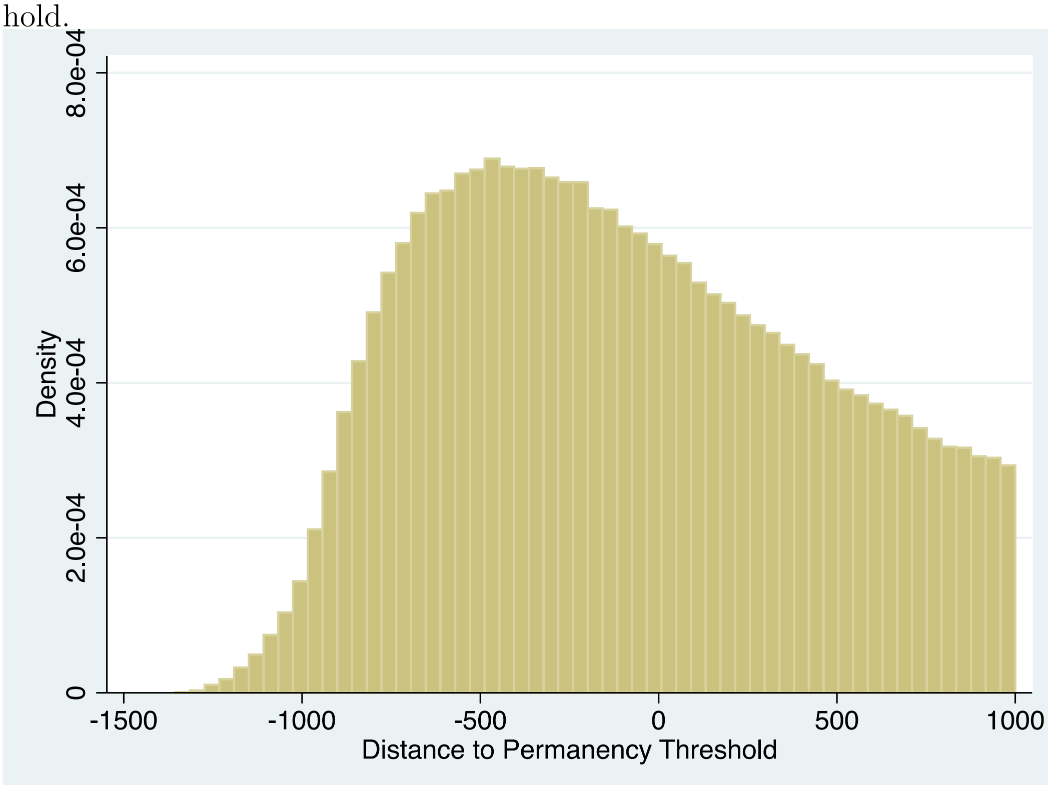

Table 1: Descriptive statistics for principal outcomes by group.

	Rural						Urban					
	Males			Females			Males			Females		
	Obs.	Mean	Std. Dev.	Obs.	Mean	Std. Dev.	Obs.	Mean	Std. Dev.	Obs.	Mean	Std. Dev.
13-15 Year Olds:												
Attends School	5,697	0.91	0.29	5,467	0.93	0.25	14,282	0.89	0.31	13,887	0.92	0.27
Work	5,687	0.07	0.25	5,457	0.02	0.14	14,256	0.09	0.28	13,871	0.03	0.16
No Work	5,687	0.06	0.23	5,457	0.03	0.17	14,256	0.05	0.22	13,871	0.04	0.20
Household Work	5,687	0.00	0.07	5,457	0.04	0.20	14,256	0.01	0.09	13,871	0.04	0.20
Studies	5,687	0.87	0.33	5,457	0.91	0.29	14,256	0.85	0.35	13,871	0.89	0.31
Indigenous	5,603	0.25	0.43	5,375	0.25	0.43	14,062	0.23	0.42	13,693	0.23	0.42
Av. Educ. Parents	5,697	7.54	2.63	5,468	7.52	2.69	14,287	7.93	2.82	13,892	7.94	2.86
No Spouse	5,697	0.18	0.38	5,468	0.19	0.39	14,287	0.21	0.41	13,892	0.22	0.41
Num. Oth. Minors	5,697	1.52	1.11	5,468	1.56	1.16	14,287	1.61	1.12	13,892	1.62	1.11
Nuclear HH	5,697	0.84	0.37	5,468	0.83	0.38	14,287	0.80	0.40	13,892	0.80	0.40
HH Size	5,697	4.80	1.42	5,468	4.87	1.52	14,287	4.93	1.52	13,892	4.92	1.52
Age Head	5,697	42.91	7.60	5,468	43.14	7.75	14,287	42.23	7.60	13,892	42.22	7.68
Agricultural Land	5,697	0.15	0.36	5,468	0.16	0.36	14,287	0.05	0.21	13,892	0.05	0.21
Food Vulnerability	5,697	0.39	0.49	5,468	0.38	0.49	14,287	0.41	0.49	13,892	0.40	0.49
HH Income	4,781	5,821	15,036	4,599	5,629	4,730	12,814	6,715	4,899	12,445	6,713	10,094
16-18 Year Olds:												
Attends School	4,346	0.69	0.46	4,060	0.75	0.44	10,857	0.69	0.46	10,435	0.74	0.44
Work	4,339	0.26	0.44	4,054	0.09	0.28	10,851	0.28	0.45	10,417	0.13	0.33
No Work	4,339	0.09	0.29	4,054	0.05	0.22	10,851	0.08	0.27	10,417	0.06	0.23
Household Work	4,339	0.00	0.07	4,054	0.15	0.36	10,851	0.01	0.10	10,417	0.12	0.32
Studies	4,339	0.64	0.48	4,054	0.71	0.45	10,851	0.63	0.48	10,417	0.70	0.46
Indigenous	4,266	0.26	0.44	4,005	0.26	0.44	10,674	0.24	0.43	10,311	0.25	0.43
Av. Educ. Parents	4,348	7.18	2.80	4,062	7.23	2.78	10,866	7.68	2.95	10,440	7.65	2.94
No Spouse	4,348	0.18	0.38	4,062	0.19	0.40	10,866	0.23	0.42	10,440	0.24	0.43
Num. Oth. Minors	4,348	1.45	1.16	4,062	1.45	1.19	10,866	1.50	1.15	10,440	1.51	1.16
Nuclear HH	4,348	0.78	0.41	4,062	0.77	0.42	10,866	0.76	0.43	10,440	0.72	0.45
HH Size	4,348	4.91	1.56	4,062	4.92	1.57	10,866	4.99	1.58	10,440	5.03	1.60
Age Head	4,348	45.97	7.72	4,062	46.00	7.76	10,866	44.95	7.65	10,440	45.05	7.65
Agricultural Land	4,348	0.19	0.40	4,062	0.18	0.39	10,866	0.06	0.24	10,440	0.06	0.23
Food Vulnerability	4,348	0.42	0.49	4,062	0.39	0.49	10,866	0.42	0.49	10,440	0.40	0.49
HH Income	3,746	6,151	17,102	3,401	5,762	5,089	9,769	6,995	5,047	9,392	6,834	5,278

Table 2: Results For School Attendance.

		All		Rural	Urban	Rural		Urban	
	All	Male	Female	All	All	Male	Female	Male	Female
13-15 Years:									
RDD Effect	-0.0377**	-0.0461**	-0.0270	0.0051	-0.0567***	0.0021	0.0096	-0.0559**	-0.0497**
p-value	(0.0238)	(0.0459)	(0.1699)	(0.8184)	(0.0050)	(0.9489)	(0.7626)	(0.0387)	(0.0489)
Obs. left	3998	2262	2461	2362	2590	1258	1030	1627	1426
Obs. right	3809	2124	2392	2328	2542	1248	996	1508	1410
Bandwidth Regression	20.08	22.36	25.08	42.28	18.32	44.24	36.74	22.33	20.33
Bandwidth Bias	33.67	37.04	43.03	62.21	31.50	65.05	54.47	35.77	36.12
16-18 Years:									
RDD Effect	-0.0371	-0.0066	-0.0495	-0.1387***	0.0003	-0.1432**	-0.1496***	0.0163	-0.0091
p-value	(0.1248)	(0.8347)	(0.1181)	(0.0022)	(0.9928)	(0.0253)	(0.0086)	(0.7087)	(0.7962)
Obs. left	4492	2764	2453	1131	3237	607	564	1558	2189
Obs. right	4448	2857	2318	1102	3209	599	554	1651	2068
Bandwidth Regression	30.12	36.75	32.90	26.73	30.25	27.87	27.79	29.40	40.76
Bandwidth Bias	52.46	58.95	55.59	49.18	47.53	47.40	52.52	46.81	62.35

Notes: Results show bias-corrected estimates for discontinuity using local linear regression; ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. Robust p-values in parentheses

Table 3: Results for Main Activity in the Past Week.

	Male				Female			
	Paid Work	No Work	HH Work	Studying	Paid Work	No Work	HH Work	Studying
13-15 Years								
Urban:								
RDD Effect	0.0213	0.0215	0.0150*	-0.0530*	0.0191	0.0059	-0.0080	-0.0299
p-value	(0.3362)	(0.3141)	(0.0538)	(0.0942)	(0.1200)	(0.7220)	(0.5823)	(0.2659)
Obs. left	1958	1754	1808	1691	2310	2220	2504	1768
Obs. right	1809	1617	1656	1571	2286	2217	2478	1744
Bandwidth Regression	26.85	23.87	24.53	23.18	33.26	32.11	36.25	25.25
Bandwidth Bias	41.48	37.83	35.48	35.39	55.91	48.37	55.04	43.59
16-18 Years								
Rural:								
RDD Effect	0.0646	0.0093	0.0158*	-0.0910	0.0391	0.0546**	0.0450	-0.1335**
p-value	(0.2959)	(0.8169)	(0.0785)	(0.1930)	(0.3188)	(0.0408)	(0.4180)	(0.0389)
Obs. left	595	640	571	574	555	442	563	573
Obs. right	588	632	560	562	544	429	552	561
Bandwidth Regression	27.25	29.37	26.09	26.35	27.32	21.90	27.74	28.29
Bandwidth Bias	44.72	45.84	30.44	41.73	45.01	37.49	46.63	46.65

Notes: Results show bias-corrected estimates for discontinuity using local linear regression; ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. Robust p-values in parentheses

Table 4: Results for Other Characteristics.

	Indigenous	Educ Parents	No Spouse	Other Minors	Nuclear HH	HH Size	Age Head	Agri Land	Food Vul	HH Income
13-15 Years										
Urban:										
RDD Effect	0.0122	-0.1814	-0.0275	-0.1184*	-0.0068	-0.0045	0.8648*	0.0016	0.0279	-624.12
p-value	(0.5750)	(0.2266)	(0.1999)	(0.0908)	(0.7335)	(0.9592)	(0.0709)	(0.8735)	(0.3535)	(0.2595)
Obs. Left	5039	5216	4873	3285	5602	4610	3645	5174	3937	2185
Obs. Right	4874	5053	4712	3144	5386	4457	3480	5011	3834	2152
BW Regression	35.89	36.53	34.18	22.92	39.19	32.32	25.39	36.26	27.82	17.31
BW Bias	56.41	60.37	56.64	40.18	59.57	48.82	41.66	55.94	43.04	38.13
16-18 Years										
Rural:										
RDD Effect	-0.0037	-0.0776	0.0053	-0.0451	-0.0224	0.0553	-0.3502	-0.0153	0.0343	49.81
p-value	(0.9258)	(0.8022)	(0.8789)	(0.7119)	(0.5610)	(0.7365)	(0.6978)	(0.6624)	(0.4566)	(0.9252)
Obs. Left	1607	1203	1648	1223	1516	1244	1009	1814	1615	1229
Obs. Right	1581	1189	1619	1202	1500	1236	988	1779	1590	1224
BW Regression	38.42	28.60	38.71	28.94	35.76	29.61	24.23	42.58	37.96	34.40
BW Bias	61.08	43.86	60.44	43.72	57.25	47.38	38.90	63.96	57.37	53.40

Notes: Results show bias-corrected estimates for discontinuity using local linear regression; ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. Robust p-values in parentheses

Table 5: Results for School Attendance at Different Polynomials and Data Windows.

	13-15 Year Olds Urban			16-18 Year Olds Rural		
	+/- 50	+/- 100	+/- 200	+/- 50	+/- 100	+/- 200
First Order						
RDD Effect	-0.0625***	-0.0567***	-0.0235*	-0.1155*	-0.1387***	-0.1199***
p-value	(0.0046)	(0.0050)	(0.0579)	(0.0595)	(0.0022)	(0.0033)
Obs. left	2151	2590	7184	669	1131	1421
Obs. right	2128	2542	6891	665	1102	1421
Bandwidth Regression	15.20	18.32	50.15	15.92	26.73	33.48
Bandwidth Bias	26.12	31.50	83.94	24.57	49.18	61.44
Second Order						
RDD Effect	-0.0690**	-0.0596***	-0.0204	-0.0950	-0.1465**	-0.1162***
p-value	(0.0101)	(0.0036)	(0.1132)	(0.2342)	(0.0100)	(0.0077)
Obs. left	2878	4963	13616	771	1514	2697
Obs. right	2763	4803	12865	741	1499	2608
Bandwidth Regression	20.15	34.80	93.91	18.20	35.72	62.29
Bandwidth Bias	27.79	52.32	128.75	24.55	48.81	86.32
Third Order						
RDD Effect	-0.0655**	-0.0632***	-0.0615***	-0.0859	-0.1448**	-0.1134**
p-value	(0.0444)	(0.0070)	(0.0016)	(0.3016)	(0.0201)	(0.0144)
Obs. left	2935	6481	9094	1152	2080	3962
Obs. right	2821	6233	8760	1134	2048	3730
Bandwidth Regression	20.56	45.29	63.59	27.31	49.09	91.58
Bandwidth Bias	26.33	57.87	86.78	34.54	61.33	114.55

Notes: Results show bias-corrected estimates for discontinuity using local polynomial regression of first to third order; ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. Robust p-values in parentheses

Figure 2: Discontinuity for school outcome with 3rd order polynomial for urban 13-15 year-olds.

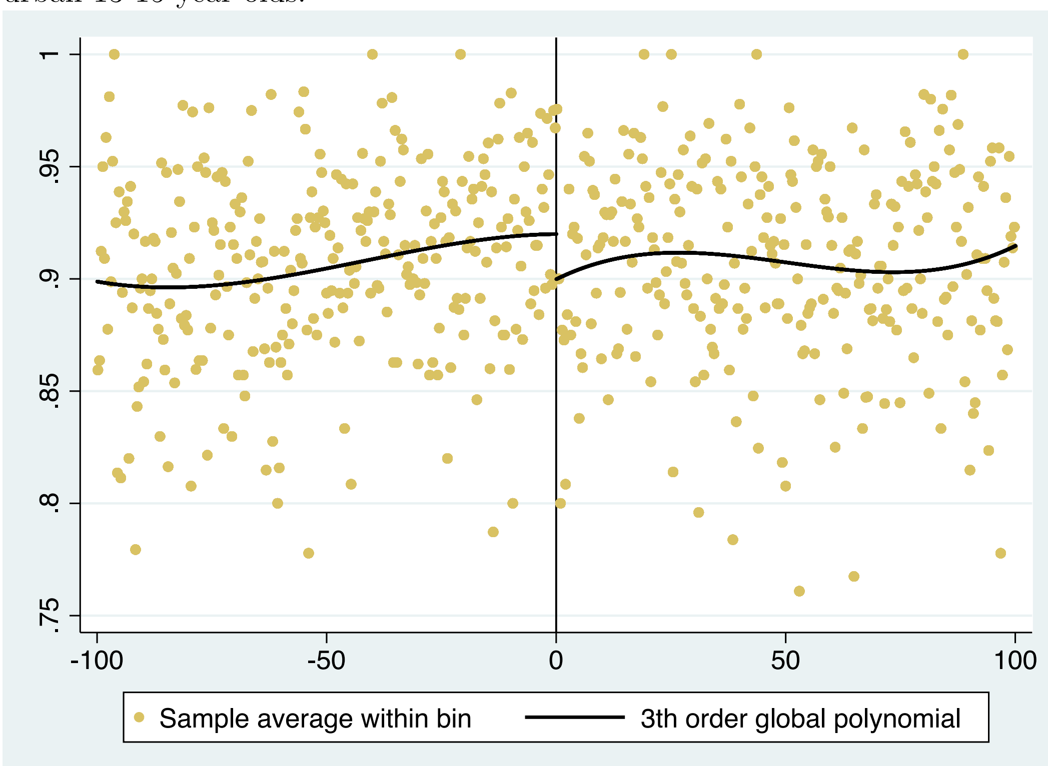

Figure 3: Discontinuity for school outcome with 3rd order polynomial for rural 16-18 year olds.

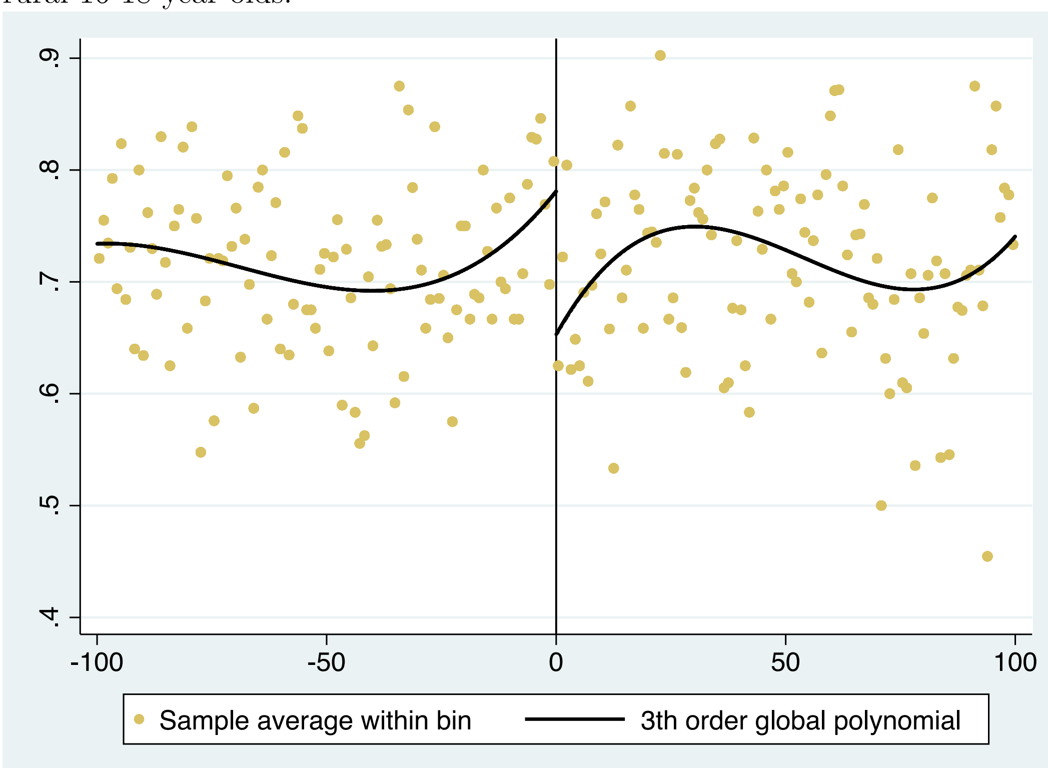

Table A.1: Results for determinants of having VHS or DVD equipment.

	Urban	Rural
Apartment	0.123**	0.910***
	(0.0555)	(0.315)
Solid Walls	0.157***	0.103**
	(0.0462)	(0.0470)
Solid Roof	0.0814**	0.0864*
	(0.0338)	(0.0453)
Kitchen	0.0960*	-0.0292
	(0.0508)	(0.0653)
Public Sewer	-0.0137	0.0846*
	(0.0324)	(0.0445)
Number Light Bulbs	0.0429***	0.0542***
	(0.00449)	(0.00871)
Shower	0.0181	-0.0607
	(0.0357)	(0.0541)
Water Tank	0.0398	0.0482
	(0.0271)	(0.0530)
Cistern	0.0805**	0.155**
	(0.0340)	(0.0745)
Boiler	0.210***	0.0642
	(0.0268)	(0.0590)
Water Pump	0.0628*	0.0619
	(0.0327)	(0.0585)
Pay TV	0.187***	0.143***
	(0.0239)	(0.0434)
Radio	-0.0949***	-0.208***
	(0.0296)	(0.0523)
Washer	0.368***	0.315***
	(0.0266)	(0.0415)
Constant	-1.149***	-1.092***
	(0.0629)	(0.0693)
Observations	14,210	5,248

Table A.2: Parameters of estimated log income per-capita model used to assess eligibility and graduation.

	Urban	Rural
Intercept	8.245	7.389
Household head and spouse: Have completed primary, but not secondary school	0.066	0.137
Household head and spouse: Have completed secondary school or higher	0.257	0.313
Household dependency ratio	-0.034	-0.06
Number of women in household aged 15-49	-.027	
Logarithm of the total number of HH members	-0.737	-0.624
Number of HH members who are employees	0.24	0.374
Number of HH members who work independently	0.172	0.101
Number of HH members who work in a employment-like, but unpaid capacity	0.06	
Number of HH members with Seguro Popular coverage	-0.009	
At least one HH member has employer provided heath insurance	0.224	0.388
HH works independently & at least one HH member has employer provided health insurance	0.055	0.219
Most of the dwelling has solid floors		0.096
Most of the dwelling has covered floors	0.135	0.302
Dwelling is owner occupied	0.035	
Dwelling is rented	0.183	0.186
Total number of rooms	0.051	0.024
Wood or coal are used as cooking fuel	-0.112	-0.271
Dwelling has exclusive water toilet	0.015	0.074
Does not have refrigerator	-0.023	-0.121
Does not have telephone landline	-0.072	
Does not have motor vehicle	-0.23	-0.197
Does not have VHS/DVD/Blue Ray player	-0.128	-0.111
Does not have computer	-0.288	
Does not have electric or microwave oven	-0.115	-0.114
Some adult HH member did not have breakfast, lunch or supper due to lack of resources (Food Security 1)	-0.1	
Some adult HH member had only one meal a day or none at all due to lack of resources (Food Security 2)	-0.058	
Food Security 1 or Food Security 2		-0.096
Receives international remittances	0.078	0.279
Municipal social gap index	-0.047	-0.071
Locality with 100,000 inhabitants or more	0.058	
Locality with 15,000-100,000 inhabitants	0.054	