



INTERNATIONAL FOOD  
POLICY RESEARCH INSTITUTE

*sustainable solutions for ending hunger and poverty*

Supported by the CGIAR

**IFPRI Discussion Paper 01116**

**September 2011**

## **Using the Regression Discontinuity Design with Implicit Partitions**

The Impacts of *Comunidades Solidarias Rurales* on  
Schooling in El Salvador

**Alan de Brauw**

**Daniel Gilligan**

**Markets, Trade and Institutions Division**

**Poverty, Health, and Nutrition Division**

## **INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE**

The International Food Policy Research Institute (IFPRI) was established in 1975. IFPRI is one of 15 agricultural research centers that receive principal funding from governments, private foundations, and international and regional organizations, most of which are members of the Consultative Group on International Agricultural Research (CGIAR).

## **PARTNERS AND CONTRIBUTORS**

IFPRI gratefully acknowledges the generous unrestricted funding from Australia, Canada, China, Denmark, Finland, France, Germany, India, Ireland, Italy, Japan, the Netherlands, Norway, the Philippines, South Africa, Sweden, Switzerland, the United Kingdom, the United States, and the World Bank.

## **AUTHORS**

**Alan de Brauw, International Food Policy Research Institute**  
Senior Research Fellow, Markets, Trade and Institutions Division  
[a.debrauw@cgiar.org](mailto:a.debrauw@cgiar.org)

**Daniel Gilligan, International Food Policy Research Institute**  
Senior Research Fellow, Poverty, Health, and Nutrition Division  
[d.gilligan@cgiar.org](mailto:d.gilligan@cgiar.org)

## **Notices**

IFPRI Discussion Papers contain preliminary material and research results. They have been peer reviewed, but have not been subject to a formal external review via IFPRI's Publications Review Committee. They are circulated in order to stimulate discussion and critical comment; any opinions expressed are those of the author(s) and do not necessarily reflect the policies or opinions of IFPRI.

Copyright 2011 International Food Policy Research Institute. All rights reserved. Sections of this material may be reproduced for personal and not-for-profit use without the express written permission of but with acknowledgment to IFPRI. To reproduce the material contained herein for profit or commercial use requires express written permission. To obtain permission, contact the Communications Division at [ifpri-copyright@cgiar.org](mailto:ifpri-copyright@cgiar.org).

## Contents

Abstract	v
Acknowledgments	vi
1. Introduction	1
2. Regression Discontinuity Design	3
3. Partitioned Cluster Analysis	5
4. Implicitly Defining the Threshold	6
5. <i>Comunidades Solidarias Rurales</i>	9
6. Data Sources	12
7. Results	16
8. Conclusion	25
Appendix: Supplementary Table	26
References	27

## List of Tables

6.1—School enrollment rates in El Salvador in 2006 and 2007, by age	13
6.2—School enrollment rates in El Salvador among children 7 to 12 years of age, in 2006 and 2007, by year of entry into <i>Comunidades Solidarias Rurales</i>	13
6.3—Percent of children 6 to 12 years of age enrolled in school in rural El Salvador in 2007, by age and gender	14
6.4—School enrollment rates in rural El Salvador in 2007, by year of entry into <i>Comunidades Solidarias Rurales</i> and by gender	14
6.5—School enrollment rates in rural El Salvador in 2007, by poverty group	14
7.1—Regression discontinuity results for impact of transfer associated with <i>Comunidades Solidarias Rurales</i> on change in school enrollment rates in El Salvador from 2006 to 2007 among children 7 to 12 years of age, comparing 2006 entrants with 2007 entrants	17
7.2—Regression discontinuity results for impact of transfer associated with <i>Comunidades Solidarias Rurales</i> on school enrollment rates in El Salvador in 2007 among children 7 to 12 years of age, comparing 2006 entrants with 2007 entrants	19
7.3—Impact of <i>Comunidades Solidarias Rurales</i> on school enrollment in El Salvador, by age and gender, bandwidth of 5	21
7.4—Regression discontinuity results for impact of transfer associated with <i>Comunidades Solidarias Rurales</i> on school enrollment rates in El Salvador in 2007 among children 6 years of age, comparing 2006 entrants with 2007 entrants	23
7.5—Regression discontinuity results for impact of transfer associated with <i>Comunidades Solidarias Rurales</i> on school enrollment rates in El Salvador in 2007 among children 6 years of age and children 7 to 12 years of age, controlling for individual and household characteristics, bandwidth of 5	24
A.1—Regression discontinuity results for impact of transfer associated with <i>Comunidades Solidarias Rurales</i> on school enrollment rates in El Salvador in 2007 among children 7 to 12 years of age, comparing 2006 entrants with 2007 entrants, using poverty rate as the forcing variable	26

## List of Figures

4.1—Illustration of implicit threshold in a two-dimensional case	7
5.1—Municipalities in El Salvador by extreme poverty group clusters with cluster centers	10
5.2—Municipality distance from implicit cluster threshold for severe and high extreme poverty group municipalities in El Salvador, 2006 and 2007 <i>Comunidades Solidarias Rurales</i> entrants	11
6.1—Relationship between difference in distance to cluster centers and other variables measured at the <i>municipio</i> level in El Salvador	15
7.1—Change in school enrollment rate in El Salvador from 2006 to 2007 among children 7 to 12 years of age by distance from implied cluster threshold, 2006 and 2007 entry groups	17
7.2—Average net school enrollment rates in El Salvador in 2007 at the <i>municipio</i> level among children 7 to 12 years of age, comparing 2006 entry group with 2007 entry group	19
7.3—Average net enrollment rates in El Salvador among children 7 to 12 years of age, by age and <i>municipio</i> level, comparing 2006 entry group with 2007 entry group	20
7.4—Smoothed impact estimates using alternative forcing variable (poverty rate), children 7 to 12 years of age in El Salvador	22
7.5.—Average net enrollment rates, 6 year olds, Municipio level, Comparing 2006 to 2007 Entry group, El Salvador	23

## ABSTRACT

Regression discontinuity design (RDD) is a useful tool for evaluating programs when a single variable is used to determine program eligibility. RDD has also been used to evaluate programs when eligibility is based on multiple variables that have been aggregated into a single index using explicit, often arbitrary, weights. In this paper, we show that under specific conditions, regression discontinuity can be used in instances when more than one variable is used to determine eligibility, without assigning explicit weights to map those variables into a single measure. The RDD approach used here groups observations that are common across multiple criteria through the use of a distance metric to create an implicit partition between groups. We apply this model to evaluate the impact of the conditional cash transfer program *Comunidades Solidarias Rurales* in El Salvador, which used partitioned cluster analysis to determine the order communities would enter the program in a phased roll-out, as a function of the poverty rate and severe stunting rate. Using data collected for the evaluation as well as data from the 6th National Census of El Salvador, we demonstrate that the program increased both *parvularia* and primary school enrollment among children aged 6 to 12 years old. Among children of primary school age, enrollment gains were largest among younger children and older girls.

**Keywords:** regression discontinuity design, partitioned cluster analysis, schooling, impact evaluation, El Salvador

## ACKNOWLEDGMENTS

We thank Mauricio Shi Artiga, Doug Miller, Amber Peterman, Margarita Beneke de Sanfeliu, Mauricio Sandoval, and seminar participants at the University of California, Davis, for contributions and suggestions that have strengthened this paper. The evaluation data used in the study were collected on behalf of the government of El Salvador through the *Fondo de Inversión Social para el Desarrollo Local* (FISDL). Please direct correspondence to Alan de Brauw at [a.debrauw@cgiar.org](mailto:a.debrauw@cgiar.org). All remaining errors are our responsibility.

# 1. INTRODUCTION

Regression discontinuity design (RDD) methods have become increasingly popular in evaluating the impacts of social programs in the economics literature. In general, evaluations have been based on applying RDD around thresholds of a single metric that determines program eligibility (see Lee and Lemieux 2010 for a review). Since the threshold is arbitrary from the perspective of the unit of intervention, units that are just eligible for the program—or have values of the metric “close” to the threshold—can be compared with units that are just not eligible, to measure the local average treatment effect of the program.

It is not necessarily the case that one metric determines program eligibility. Rather, governments or agencies charged with determining program eligibility may instead choose to use two or more metrics. If two or more metrics are used, a common way to map these into a single measure is through a mathematical function that assigns explicit weights to each metric to construct the aggregate single metric. If this procedure is used, then regression discontinuity is still simple to use if specific assumptions are met (see, for example, Imbens and Lemieux 2008). A common example of such a procedure is when program eligibility is determined by a proxy means test, which effectively turns several measures into one measure that can then be used to determine strict program eligibility (see, for example, Galasso 2006; Chaudhury and Parajuli 2006; Filmer and Schady 2009; and Ponce and Bedi 2010).

However, one does not necessarily need to use a well-defined function to determine program eligibility. For example, Kane (2003) considers the case in which students graduating from high school in California only become eligible for grants if they achieved a minimum grade-point average (GPA) and had income and financial assets below specific thresholds. In several other cases, authors use the distance to a city boundary as a forcing variable, which technically depends upon two different variables for a threshold, longitude and latitude (Black 1999; Bayer, Ferreira, and McMillan 2007). Anti-poverty programs, however, are often targeted on the basis of a proxy means test or a similar function that transforms several variables into a single number. One can argue that such functions use arbitrary or politically determined weights assigned by program managers to determine the score. The approach to RDD proposed here uses a less arbitrary method of grouping observations that are similar across several characteristics. In this approach, partitioned cluster analysis is used to identify similar groups within data, which classifies individual observations into similar clusters of observations. If a subset of those clusters is then assigned a treatment, then the treatment status is completely determined by cluster membership and therefore by the metrics used in assigning units to clusters. But an explicit threshold between treatment clusters and control clusters does not exist, so one cannot immediately perform regression discontinuity to determine program impacts.

In this paper, we develop a set of additional assumptions needed to use standard regression discontinuity methods to evaluate programs that determine treatment status using partitioned cluster analysis or similar methods. The idea behind the estimator is that we use the distance metric that determines clustering in the data to implicitly define the threshold between treatment and control groups as a function of the distance between cluster centers. Under quite reasonable assumptions, we show that the threshold can then be used in a sharp regression discontinuity estimator using the distance from each point to the threshold in estimation. This approach has the attraction of relying on the data to group units that are similar across multiple metrics used to determine program eligibility rather than using arbitrary weights to aggregate across metrics. We are aware of only one other paper that has proposed a similarly general approach to using regression discontinuity to evaluate program impacts when eligibility is based on multiple criteria (Papay, Willett and Murnane 2011).

We then apply this methodology to evaluate a specific conditional cash transfer program—Comunidades Solidarias Rurales (CSR) in El Salvador—that used partitioned cluster analysis to determine the order in which municipios (municipalities) entered the program as well as which municipios would receive CSR benefits. Using data from the evaluation of CSR as well as data from the sixth National Census of El Salvador (Censo Nacional de Poblacion y Vivienda de 2007), we compare schooling outcomes among households in municipios that entered CSR in 2006 and households in municipios that entered in 2007. We find that close to the threshold, children in the 2006 entry group were more likely to have

enrolled in parvularia (preschool) at age 6, and that school enrollment rates among primary-school age children increased by 4 percentage points. Using the census data, we further disaggregate these results by age and gender.

The paper is organized as follows: Section 2 briefly reviews the one-dimensional regression discontinuity estimator, including assumptions necessary for the estimator to provide an unbiased estimate of the treatment effect; Section 3 provides a brief description of partitioned cluster analysis; Section 4 develops conditions for an n-dimensional regression discontinuity estimator to be valid; Section 5 presents basic information about CSR; Section 6 describes the data sources used for analysis; Section 7 presents results; and Section 8 concludes.



## 2. REGRESSION DISCONTINUITY DESIGN

Regression discontinuity designs are typically referred to as being either sharp or fuzzy. The estimator we will develop follows the sharp design, so we review it here. Following the notation of Imbens and Wooldridge (2009), we can consider two potential outcomes for unit  $i$ , namely  $Y_i(0)$  and  $Y_i(1)$ , where the difference  $Y_i(1) - Y_i(0)$  is the definition of the causal effect of the treatment. The observed outcome is equal to

$$Y_i = (1 - W_i) \cdot Y_i(0) + W_i \cdot Y_i(1), \quad (1)$$

where  $W_i \in \{0,1\}$  is the treatment indicator variable. The idea behind a sharp regression discontinuity evaluation is that there is a variable  $X_i$  that completely determines whether or not a unit receives the treatment. Calling this threshold  $c$ , a unit will receive the treatment if  $X_i \geq c$ , which implies

$$W_i = 1\{X_i \geq c\}. \quad (2)$$

In a sharp regression discontinuity design, all units with a value of  $X_i$  that is at least  $c$  do receive the treatment, and those units with a value of  $X_i$  below  $c$  do not receive the treatment, effectively becoming the control group. The average treatment effect  $\delta$  is the difference between the mean outcome for units with values of  $X_i$  just below the threshold ( $Y^-$ ) and those with values of  $X_i$  just above the threshold ( $Y^+$ ). It can then be written as the difference in conditional expectations between units just above and just below the threshold:

$$\delta = Y^+ - Y^- = \lim_{\epsilon \rightarrow 0} E(Y_i(1)|X_i = c + \epsilon) - E(Y_i(0)|X_i = c - \epsilon) \quad (3)$$

To estimate  $\delta$ , one needs to estimate both ( $Y^+$ ) and ( $Y^-$ ). One can generally write the solution to the estimation problem in the form of nonparametric regressions:

$$\hat{Y}^+ = \frac{\sum_i X_i > c Y_i K\left(\frac{X_i - c}{h}\right)}{\sum_i X_i \geq c Y_i K\left(\frac{X_i - c}{h}\right)} \quad (4)$$

$$\hat{Y}^- = \frac{\sum_i X_i < c Y_i K\left(\frac{c - X_i}{h}\right)}{\sum_i X_i \geq c K\left(\frac{c - X_i}{h}\right)} \quad (5)$$

where  $K(\cdot)$  represents a kernel estimator and  $h$  represents the chosen bandwidth. The major problem here is to choose a kernel function that will identify the effect at the single point of interest, the threshold, as well as a proper bandwidth. Porter (2003) shows that the bias in the estimate using the rectangular kernel is linear in the bandwidth  $h$ , whereas the bias in nonparametric estimators generally is of order  $h^2$  in nonparametric estimators. In this paper, we vary the bandwidth to test the sensitivity of results to inclusion or exclusion of observations farther away from the threshold.<sup>1</sup>

Three assumptions are critical for the consistency of the sharp regression discontinuity estimator (Edmonds, Mammen, and Miller 2005). First, the probability of treatment must vary discontinuously at the threshold. Intuitively, the sharp cutoff point serves as an instrumental variable that affects program participation but does not independently affect outcomes. Second, observations just above and just below the threshold must be similar in both their observed and unobserved characteristics. Third, one must assume

---

<sup>1</sup> See Ludwig and Miller (2007) and Imbens and Lemieux (2008) for details on methods of choosing the bandwidth in sharp regression discontinuity applications. We do not apply these methodologies here because we use a community-level forcing variable rather than an individual-level forcing variable.

that if the treatment did not occur, the outcome  $Y_i$  would be continuous at the threshold. In other words, without the program, there would be no sharp break in outcome measures in the population at large between those just below and those just above the threshold. Another way to think of the RDD estimator is as generating a locally randomized experiment (Lee and Lemieux 2010), because if the continuity assumption is met, then the forcing variable takes on almost the same value at the threshold.

### 3. PARTITIONED CLUSTER ANALYSIS

Partitioned cluster analysis encompasses a set of iterative methods of breaking observations in datasets into distinct groups that are similar across multiple characteristics. Observations are grouped along those attributes using an iterative procedure that proceeds as follows. First, the analyst chooses the number of groups,  $K$ , he/she is interested in forming. Second, the analyst chooses  $K$  initial points (centers), calculates the distance from each observation to each initial point, and assigns each observation the center nearest to it. The centers of each cluster are recalculated as either the mean or median of the points aligned with that center, and distances to the new  $K$  centers are computed. Clusters are then re-formed if observations switch groups, and the procedure is repeated until all points no longer switch clusters.<sup>2</sup> We define the center of the  $k$ -th cluster as  $\eta_k$  and the distance between any observation  $X_i$  and  $\eta_k$  as  $d(X_i, \eta_k)$ . Cluster membership is then defined for all clusters  $k$  as the set of all  $X_{ik}$ , such that

$$X_{ik} = \arg \min_{k \in K} (d(X_i, \eta_k)). \quad (6)$$

The analyst must initially determine how many clusters to use. The number of clusters can be determined through data driven processes; for example, in cluster  $k$ -means analysis the number of clusters that maximize the pseudo- $F$  statistic of Calinski and Harabasz (1974) can be chosen. To set up the estimator based on the results of cluster analysis, we assume that the analyst has chosen a finite number of clusters  $K$ ; the initial centers of each cluster; the distance measure to be used; and the method of choosing new cluster centers (mean or median). After these choices are made, the results of the cluster analysis will always be the same; that is, they are always replicable with the same data source. Furthermore, important from the perspective of performing a regression discontinuity analysis on partition clustered data, each point distinctly belongs to one cluster. Therefore the estimation strategy suggested is similar to that in a sharp regression discontinuity design.

---

<sup>2</sup> When the mean is used as the center, partitioned cluster analysis is usually known as cluster  $k$ -means analysis; when it is the median, it is known as cluster  $k$ -medians analysis.

## 4. IMPLICITLY DEFINING THE THRESHOLD

When program assignment is performed using partitioned cluster analysis, then there is no true forcing variable. In this section, we describe additional assumptions under which an implicit forcing variable can be defined when partitioned cluster analysis is used with  $M$  traits.

To set up the estimator, assume a sample of  $n$  individuals indexed by  $i$ , who all have an  $M$  member vector of traits,  $X_i$ . The traits are to be used to determine which individuals receive the treatment and which individuals do not receive the treatment. Assume that each element of  $X$  is positive (for example,  $X_j \geq 0 \forall j \in M$ ). A partition cluster analysis is performed on these individuals to create  $k$  clusters, and each individual can then be indexed as  $X_{ik}$ . Individuals are then partitioned into clusters, and the analyst chooses  $a$  of the clusters to receive the intervention (treatment clusters) and the remaining  $b = k - a$  clusters remain as control clusters.

Given that each cluster has a well-defined center, the centers implicitly define boundaries between all  $k$  of the clusters. These boundaries can be defined as follows. Consider any two cluster centers,  $\eta_c$  and  $\eta_d$ . There must be a set of points that are equidistant from the two cluster centers,  $W_j$ , which define a boundary between those two cluster centers,  $d(W_j, \eta_c) = d(W_j, \eta_d)$ . This logic can extend to all  $\frac{K(K-1)}{2}$  of the clusters; between any two clusters, there must be a boundary defined as the set of points that are equidistant from the two cluster centers.

Since boundaries exist between all clusters, it must be that a boundary exists between any adjacent treatment and control clusters. Furthermore, we can also define a further boundary that relates the closest center of a treatment cluster to the closest center of a control cluster. This boundary separates treatment clusters from control clusters. If we define this set of points as  $Z_j$ , the set of points acts as the threshold between treatment and control observations. It can more formally be defined as

$$\min_{a \in A} d(Z_j, \eta_a) = \min_{a \in A} d(Z_j, \eta_b). \quad (7)$$

As long as equation (7) conforms to three specific assumptions, the implicit function theorem states that it defines a function, and so with the choice of a distance metric, that function defines a specific implicit forcing variable that can be used to perform RDD estimation.<sup>3</sup>

The first assumption is that the function defined by equation (7) must be continuous. If it is not continuous, then the control group for either certain treatment clusters or specific treatment observations might not be well defined. Under certain conditions—for example, if an analyst was focusing on one specific treatment cluster among many—then local continuity would suffice for all points local to that specific treatment cluster and its center.

Second, the solution to equation (7) must be unique; there can be only one vector  $Z_j$  that satisfies equation (7). We must make two assumptions to ensure that a unique solution holds: (1) if a unique solution does not exist, then two or more potential boundaries exist, and it is not possible to ascertain which of the boundaries would correctly define the forcing variable; and (2) from a mathematical perspective, if uniqueness did not exist, then the implicit function theorem would break down for equation (7).

Third, we make the assumption that all of the indicators used in partitioned cluster analysis increase the likelihood of being in the treatment. Formally, assume that  $X_t$  is a point in a treatment cluster. Then for any point  $X_j$ , if  $\chi_{js} \geq \chi_{ks}$  for  $s = 1, \dots, M$ , then it must be that  $X_j$  is also in the treatment. This assumption simply implies that the treatment clusters are all in generally the same neighborhood, as are the control clusters. It rules out choosing a cluster that is quite dissimilar from other clusters as a control cluster.

Under these three assumptions, the solution to equation (7) implicitly defines a boundary in one dimension—distance—between the treatment and control clusters, and as a result one can perform regression discontinuity using the distance from that set of points as the forcing variable. This estimator is

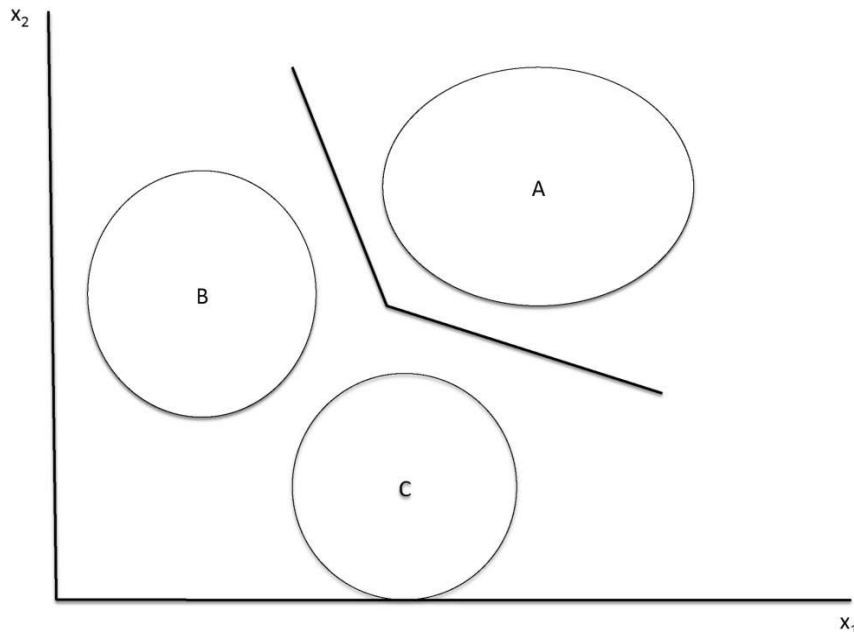
---

<sup>3</sup> The clear choice of a distance metric is the same distance metric used to perform partitioned cluster analysis.

therefore similar to those employed by authors who use the spatial distance to a city or geographic boundary as a forcing variable.<sup>4</sup> However, there are three important differences. First, this boundary need not be restricted to two dimensions. Second, the distance metric need not be Euclidean; it could, for instance, be an absolute value or a minimum distance to the boundary, which corresponds to the strategy used by Kane (2003). Third, and perhaps most important, because the boundary is implicit, one need not be concerned that individuals might purposely move across the boundary to receive a different set of services, as with city boundaries.

To further illustrate this concept, consider the two-dimensional example in Figure 4.1. The treatment cluster must be *A* by the third assumption, whereas there are two control clusters (*B* and *C*). The boundary between the treatment and control clusters is the set of points that are equidistant between the centers *A* and *B* when the center of *B* is closer, and then between *A* and *C* when the center of *C* becomes closer.

**Figure 4.1—Illustration of implicit threshold in a two-dimensional case**



Source: Authors.

To finalize the estimator, given that the set of points  $Z_j$  defines the boundary between the treatment and control groups, the distance from an individual's traits  $X_{ik}$  to  $Z_j$  defines how close the individual is to the boundary. We can define this distance  $\delta$  as

$$\delta(X_{ik}) = |X_{ik} - Z_j|. \quad (8)$$

Given the process of partitioned cluster analysis described in the previous section, points that are arbitrarily close to the boundary would be the ones where, with a small change in one of the indicators making up  $X_{ik}$ , individual  $i$  might switch from a treatment to a control cluster, or vice versa. Therefore, the distance metric  $\delta$  acts as the forcing variable in regression discontinuity. In parallel with a more explicit forcing variable, one can imagine that small shifts in the values of specific components of the cluster

<sup>4</sup> Such papers include Black (1999), Lavy (2006), Pence (2006), and Lalive (2008).

analysis would shift these observations closer to the threshold from one cluster to another, whereas those closer to the cluster centers would be less similar.

For regression discontinuity estimates using  $\delta(\cdot)$  as the forcing variable, two of the three main assumptions that are critical for the consistency of the sharp regression discontinuity estimator still apply. The first assumption—that the probability of treatment varies discontinuously at the threshold—clearly applies. However, it is still necessary that observations just above and just below the threshold be similar in both observed and unobserved characteristics, and one must assume that any outcome  $Y$  would be continuous at the threshold in the absence of a program. If the assumptions needed to define the implicit distance metric and these latter two assumptions are true, then this estimator can be used to identify program impacts using regression discontinuity.

## 5. COMUNIDADES SOLIDARIAS RURALES

El Salvador started the *Comunidades Solidarias Rurales* (CSR) conditional cash transfer program (previously called *Red Solidaria*) in 2005 to begin to alleviate poverty in the rural areas of its poorest *municipios*. The program is targeted both categorically and geographically. From the perspective of the evaluation, we consider two categories of households as beneficiaries.<sup>5</sup> Households with children 5 years of age and under were targeted for a health transfer meant to improve health and nutrition among children in that age bracket. Households with children between the ages of 6 and 15 who had not completed primary school were targeted for an education transfer. Households received the education transfer if all children between the ages of 6 and 15 were enrolled in school and attended more than 80 percent of the time each month. Households that received either the education transfer or the health transfer received US\$15 per month, and households that were eligible for both transfers received US\$20 per month.

To select the *municipios* that would participate in CSR, geographic targeting took place in two steps. First, *municipios* were grouped into four extreme poverty groups (EPGs) using partitioned cluster analysis. The two criteria that were used were the severe poverty rate, measured using representative data collected at the municipality level between 2001 and 2004, and the prevalence of severe stunting (height-for-age z-scores below -3) among first graders, measured in a census of first graders in 2000. The two measures were deemed to be alternative measures of poverty that are quite uncorrelated with one another and therefore measure different dimensions of poverty. The first two EPGs, *municipios* rated in severe and high extreme poverty, were deemed eligible for CSR, and the poorest *municipios* entered the program first.<sup>6</sup> Within each EPG, *municipios* were then ranked by a municipality marginality index (IIMM in Spanish), which is a declining index of welfare based on poverty, education levels, and housing conditions. The IIMM was used to prioritize *municipios* for entry within each EPG.

The *municipios* in the severe EPG entered CSR in 2005 and 2006, and the *municipios* in the high EPG entered between 2007 and 2009. To identify program impacts after one year, one can compare individuals residing in a set of *municipios* that entered in one year with individuals residing in *municipios* that entered the following year. This paper focuses on the difference in outcomes among individuals in the 2006 entry group and individuals in the 2007 entry group. The comparison works for educational outcomes as follows. Households beginning to receive transfers in 2006 began to receive them relatively late in the year, long after school enrollment decisions had been made.<sup>7</sup> The education transfers became conditional at the beginning of the 2007 school year, so we expect the largest impact of the program to take place in 2007 for the 2006 entry group. Similarly, the 2007 entry group did not make decisions regarding school enrollment that were conditional until the beginning of 2008.<sup>8</sup>

### Identification of Benefits Using RDD

To identify the benefits of CSR using RDD, we use children in *municipios* entering in 2006 as the treatment group and children entering in 2007 as the control group. We use this comparison for two reasons. First, RDD identifies local average treatment effects rather than average treatment effects. We are most interested in how a conditional cash transfer program affects the poor, so it is sensible to look for benefits among the poorest group possible. Since the evaluation of CSR began too late to credibly measure school enrollment

---

<sup>5</sup> A pension program for the elderly was added in 2009 for households in the poorest *municipios*; this pension does not affect the results in this paper, since the data used precede the pension program.

<sup>6</sup> Initially, the program was designed to last only three years, but in practice even the initial *municipios* entering in 2005 still receive CSR transfers and programs.

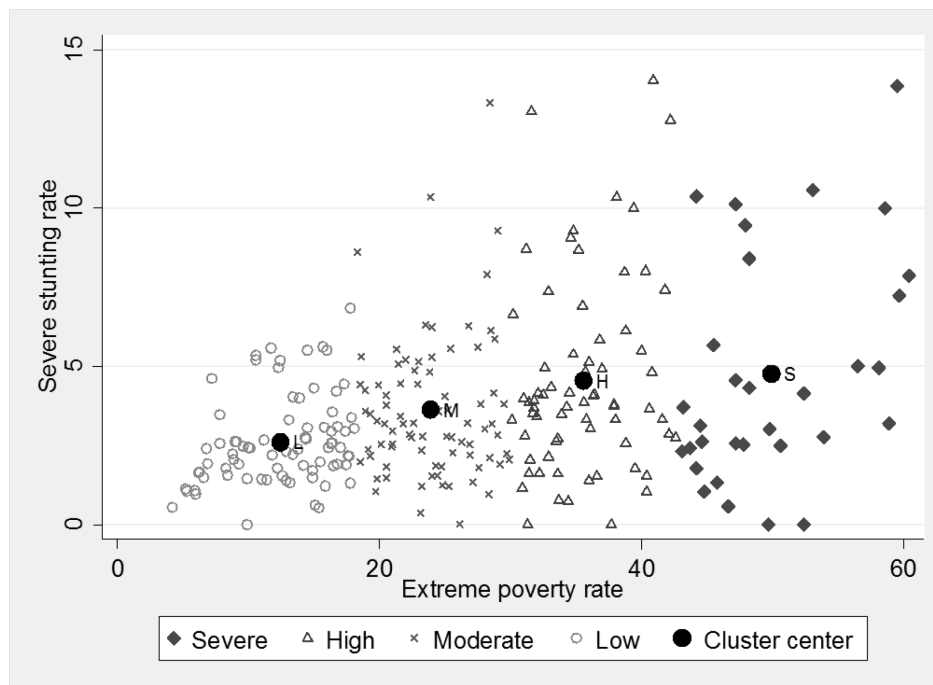
<sup>7</sup> The school year runs from January to October in El Salvador. The first CSR transfers in 2006 took place in late July in one *municipio*, and in some *municipios* the first payment did not occur until late November. As a result, program officials did not condition the education transfer until the beginning of the 2007 school year.

<sup>8</sup> One possible concern would be that people entering the program in 2007 enrolled children in school on the margin in anticipation of receiving transfers later in the year. However, more than two-thirds of families in the 2007 entry group reported learning about CSR through the program census, which took place after enrollment decisions were made, and further, any anticipation effects would make impact estimates presented here lower bounds, since they would bias results toward zero.

in the poorest group of *municipios* (those entering in 2005), we chose the 2006 entry group as the poorest possible group among which to measure impacts. Second, the data available on educational outcomes is perhaps the richest in 2007 because we can work with the El Salvador population census, which took place in May 2007 (El Salvador, Ministry of Economy 2007).

Since the 2006 entry group was in the severe EPG and the 2007 entry group was in the high EPG, the *municipios* are separated by partitioned cluster analysis rather than by IIMM ranking. As a result, we use the estimation strategy described in the previous section to develop an implicit partition between the two groups. The *municipios* in the 2006 and 2007 entry groups are illustrated in Figure 5.1. When we draw a line splitting the two clusters that is equidistant from the two cluster centers, the implicit boundary depicted conforms to the three assumptions in Section 4, and the forcing variable can be measured as the difference in distance between the cluster centers.<sup>9</sup>

**Figure 5.1—Municipalities in El Salvador by extreme poverty group clusters with cluster centers**



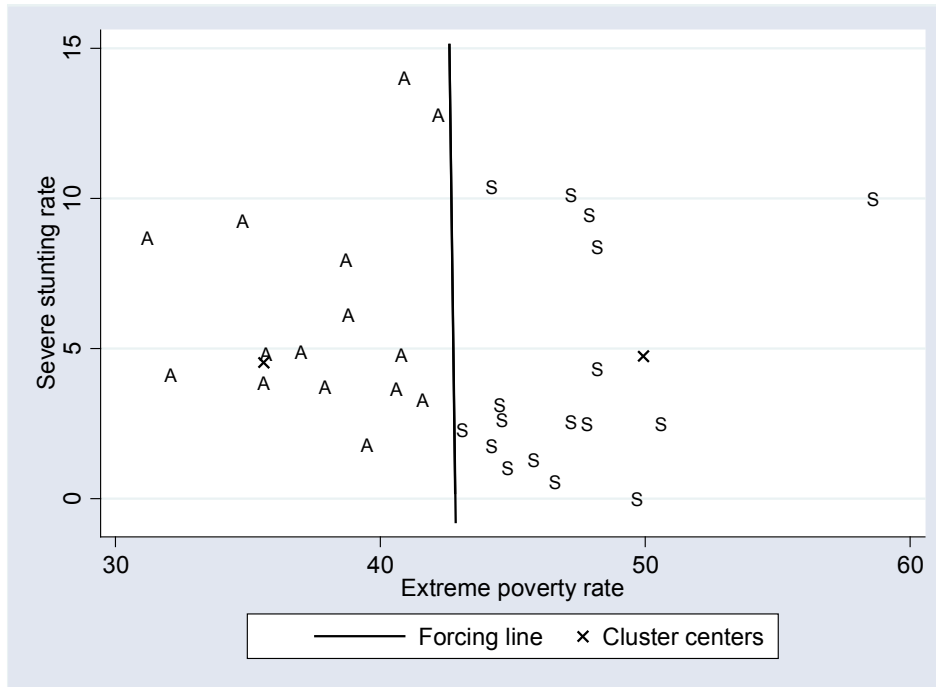
Source: Authors calculations on data from El Salvador, FISDL, 2004.

Moreover, Figure 5.2 suggests a potential robustness check. The line defined by all points equidistant from the two cluster centers is nearly vertical, and all of the *municipios* entering in 2006 have higher poverty rates than *municipios* entering in 2007. As a result, we can use the poverty rate as an alternate forcing variable, searching over poverty rates between 42.2 percent (the highest poverty rate in the 2007 group) and 43.1 percent (the lowest poverty rate in the 2006 entry group) as potential thresholds. The comparison between results using the poverty rate as the threshold and the implicit threshold gives direct evidence regarding how well the implicit threshold works as a forcing variable, since the poverty rate is simply a one-dimensional forcing variable.

<sup>9</sup> Clearly, the forcing variable is measured here as a community-level variable rather than an individual-level variable. Other papers that use community-level treatments in a regression discontinuity framework to study program impacts include Ludwig and Miller (2007), Battistin et al. (2009), and Buddelmeyer and Skoufias (2003).



**Figure 5.2—Municipality distance from implicit cluster threshold for severe and high extreme poverty group municipalities in El Salvador, 2006 and 2007 *Comunidades Solidarias Rurales* entrants**



Source: Authors calculations on data from El Salvador, FISDL, 2004.

Note: S denotes severe extreme poverty. A denotes high extreme poverty. Line indicates points equidistant from the two cluster centers.

Because of the rolling entry of the program, we can only evaluate impacts after one year of receiving transfers.<sup>10</sup> However, given that the main educational outcome we can observe is school enrollment and that the benefits of the program were conditional on school enrollment, it is reasonable to think that benefits defined as additional school enrollment were immediate. We plan to study school enrollment mainly among two groups. First, we examine children who were between the ages of 7 and 12 for the purposes of enrollment; by that, we mean that they were between 7 and 12 years old at the beginning of the calendar year. By law, children in El Salvador are supposed to be 7 years old before entering first grade, so a child progressing normally would enter his or her final year of primary school at age 12. Second, in the census data, we examine school enrollment among children 6 years of age, who are also required to be in a preschool (*parvularia*) to receive CSR transfers.

<sup>10</sup> Behrman and King (2009) show that the timing of evaluations can affect findings related to program impacts; however, due to the sequential nature of CSR entry it is not possible to measure benefits after two years of implementation, as in many impact evaluations.

## 6. DATA SOURCES

We used two data sources to evaluate the impact of CSR on school enrollment among 6- to 12-year-old children in El Salvador. The primary dataset we used was collected by the *Fundación Salvadoreña para el Desarrollo Económico y Social* (FUSADES) in collaboration with researchers at the International Food Policy Research Institute (IFPRI), and the sample was explicitly designed to include households appropriate for evaluating the impact of CSR on several indicators of infant and maternal health, education, and nutritional status, including some of the indicators used in this paper. The baseline dataset was collected in January and February of 2008 and included retrospective measures of school enrollment in a specially designed education module so that measures of school enrollment prior to CSR entry could be constructed. The survey form also included sections on household demographics, health, time allocation and off-farm labor, housing and other consumer durables, agriculture, migration, other income sources, consumer expenditures, and community participation in programs, including CSR.

The sample included 100 cantons in 50 *municipios*, distributed among the 2006 to 2008 entry groups. Prior to the baseline, 15 households with children under 3 years of age or with a pregnant woman resident, and 15 households with children between the ages of 6 and 12 were selected randomly within each canton from census lists, for a total of 30 targeted households per canton.<sup>11</sup> In some cantons, fewer households were actually interviewed. For the purposes of this paper, the most important aspect of this sample is that it included 10 *municipios* that entered CSR in 2006 and 11 *municipios* that entered in 2007, with a total of 1,280 households between those *municipios*; the entire dataset included 2,775 rural households. We focus on children in those households in the impact estimation, though we include children in the 2008 entry groups in descriptive statistics for comparative purposes.

The baseline survey specifically asked about each child's enrollment in school in 2005, 2006, and 2007; because transfers only became conditional for the 2006 entry group at the beginning of 2007, we can use data on school enrollment in 2006 and 2007 to estimate impacts in a difference-in-difference framework. Although one might be concerned about recall bias in the school enrollment measures from 2006, it should be noted that there is no reason to believe that recall bias should differ on average between the 2006 and 2007 entry groups, so estimates will not suffer from recall bias. We use the impact evaluation survey data primarily to study impacts among children aged 7 to 12 years.

Because the school year begins at the beginning of the calendar year in El Salvador and the conditionality also only begins at the beginning of the year, we can also use the national census that took place in May 2007 to study enrollment rates among all children in *municipios* entering CSR in 2006 or 2007. We created a dataset containing all of the children living in rural parts of the 2006 and 2007 entry groups to provide alternative estimates of impact among the 2006 entry group. The main drawback is that the estimates based on the census are only single difference, but since they are based on the whole rural population of relevant *municipios*, we can estimate impacts among much smaller demographic groups, for example, among age- and gender-disaggregated groups. We also estimated impacts among children 7 to 12 years of age and among 6-year-olds using the census data.

### Descriptive Statistics and Enrollment

To begin to consider how school enrollment may have changed over time as a consequence of CSR entry, we initially measured the proportion of children of each age between the ages of 7 and 12 who were enrolled in school in 2006 and 2007 (Table 6.1). It is worth noting that with the lone exception of 7-year-olds, across the entire impact evaluation sample more than 90 percent of children were reported as having been enrolled in school in 2006. In 2007, the percentages increased among all ages, by between 0.7 and 2.9 percentage points. However, given that the 2006 figures were collected through recall, it could be that there is some recall bias in the 2006 figures.

---

<sup>11</sup> The census lists were gathered by FISDL to determine which households would be eligible for benefits and which would not. The evaluation team was granted access to the relevant databases to generate a sample for the evaluation.

**Table 6.1—School enrollment rates in El Salvador in 2006 and 2007, by age**

Age	2006		2007	
	Percent enrolled	Number of observations	Percent enrolled	Number of observations
7	89.9	614	92.7	628
8	93.9	657	95.8	614
9	96.3	614	97.0	657
10	95.4	542	98.1	614
11	93.6	469	96.3	542
12	91.0	436	92.1	469

Source: Impact Evaluation Baseline Survey, *Comunidades Solidarias Rurales*, 2008.

That said, when we disaggregate 2006 and 2007 enrollment by entry group instead of by age, we find that almost all children in the 2006 entry group were reported as being enrolled in school in the impact evaluation surveys (Table 6.2). In the 2006 entry group, 98.7 percent of children were reported as being enrolled, versus between 94.2 and 95.4 percent among other entry groups. Reported enrollment increased from 2006 to 2007 in all four entry groups, though it was highest in the 2006 entry group. Changes in average enrollment rates by entry group, then, are not terribly suggestive of impacts. However, it is worth noting that there cannot be much heterogeneity in enrollment rates for the 2006 entry group, since virtually all children were enrolled in school in 2007.

**Table 6.2—School enrollment rates in El Salvador among children 7 to 12 years of age, in 2006 and 2007, by year of entry into *Comunidades Solidarias Rurales***

	Year of entry into <i>Comunidades Solidarias Rurales</i>			
	2006	2007	early 2008	late 2008
Enrolled in 2006	0.963 (0.008)	0.931 (0.014)	0.923 (0.012)	0.934 (0.013)
Enrolled in 2007	0.987 (0.004)	0.949 (0.011)	0.942 (0.009)	0.954 (0.010)

Source: Impact Evaluation Baseline Survey, *Comunidades Solidarias Rurales*, 2008.

Notes: Standard errors in parentheses clustered at the canton level. The averages for 2006 are reweighted to reflect the demographic composition of the children enrolled in 2007.

Since the census included all *municipios* in El Salvador, we can also consider enrollment by age and gender for all of rural El Salvador. The census included enough data to support examining enrollment rates by both age and gender (Table 6.3). Among all children in rural El Salvador 6 to 12 years old, we find similar enrollment rates to those described using the impact evaluation data, albeit slightly lower. Enrollment rates are fairly similar among boys and girls; girls are slightly more likely than boys to enroll in school at all ages. Enrollment rates also increased from 6-year-olds (74 percent of whom enrolled) to 10-year-olds (95 percent of whom enrolled), and then dropped off slightly. By and large, however, this table suggests that primary-school enrollment is generally quite high, as suggested by the evaluation data, and particularly among children 9 to 12 years of age.

**Table 6.3—Percent of children 6 to 12 years of age enrolled in school in rural El Salvador in 2007, by age and gender**

Age	All children	Boys	Girls
6	74.2%	73.2%	75.2%
7	88.0%	87.3%	88.6%
8	92.3%	91.8%	92.8%
9	93.8%	93.3%	94.3%
10	95.0%	94.6%	95.4%
11	94.9%	94.7%	95.1%
12	94.6%	94.5%	94.7%

Source: El Salvador, Ministry of Economy 2007.

From the perspective of the census, an open question is whether or not children in poorer areas are less likely to enroll in school after the poorest *municipios* have entered CSR. Therefore we next compute enrollment rates in 2007 by gender among the 2005, 2006, and 2007 CSR entry groups (Table 6.4). Recall that enrollment was conditional among children in the 2005 and 2006 entry groups if their families were to receive the cash transfer, whereas in the 2007 entry group, the enrollment condition did not apply until January 2008. Among 6-year-olds, the difference between the two earlier entry groups and the later group is striking. According to the census, 6-year-olds in the 2005 and 2006 entry groups were 15 percentage points more likely to be enrolled in school than those in the 2007 entry group. This finding is particularly interesting since preschool attendance was required to receive the transfer among 6-year-olds in the earlier two entry groups, and enrollment rates were much lower among children who did not reside in the severe EPG. The difference is also large among children 7 to 12 years of age, though not quite as large; the raw difference is closer to 7 percentage points between the groups receiving transfers and the group not receiving transfers. Again, we find no qualitative differences between boys and girls. These figures are strongly indicative of measurable impacts on school enrollment; we will therefore add to our impact estimation section some information from the census.

**Table 6.4—School enrollment rates in rural El Salvador in 2007, by year of entry into *Comunidades Solidarias Rurales* and by gender**

Entry year	All children		Boys		Girls	
	6	7 to 12	6	7 to 12	6	7 to 12
2005	80.5%	94.7%	81.6%	94.5%	81.3%	95.0%
2006	82.8%	95.4%	83.6%	95.3%	83.6%	95.5%
2007	65.9%	88.5%	67.2%	88.0%	67.2%	88.8%

Source: El Salvador, Ministry of Economy 2007.

In fact, the census figures are indicative of impacts in one other way. When we examine school enrollment rates among all 6- to 12-year-olds by EPG, we find that children in the severe EPG are actually more likely to be enrolled in school than children in the rest of rural El Salvador, including those in the moderate and low EPGs (Table 6.5). Given that we would expect lower enrollment rates among the poorer *municipios*, it seems likely that CSR has induced some additional enrollment.

**Table 6.5—School enrollment rates in rural El Salvador in 2007, by poverty group**

Poverty group	Enrollment rate, 6- to 12-year-olds	Number of observations
Severe	92.3%	33,183
High	86.7%	111,767
All other rural areas	89.4%	550,747

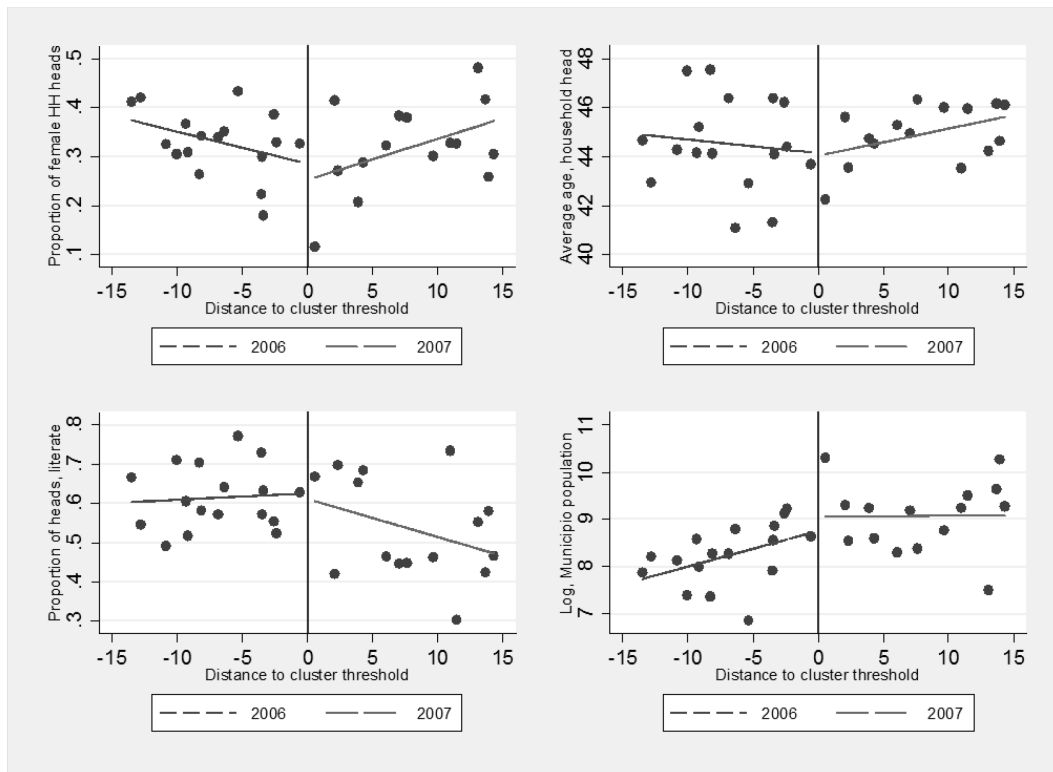
Source: El Salvador, Ministry of Economy 2007.

In summary, we find that although school enrollment rates among children of age to be in primary school are relatively high, evidence from the census conducted in 2007 is suggestive of impacts of CSR on school enrollment. Furthermore, we do not have to rely on sample estimates of enrollment to make statements about enrollment, since the census figures are the best measures of enrollment rates available for rural El Salvador. We will use the impact evaluation data and census data to construct alternative estimates of impact in the next section.

### Is the Forcing Variable Related to Average Household Characteristics?

One potential concern is that household characteristics might vary systematically near the threshold, which would invalidate results based on regression discontinuity. To visually examine whether or not the boundary is related to differing household characteristics by municipio, we graph average household characteristics against the difference in distance between cluster centers in the population census data (Figure 6.1). The characteristics we use are the proportion of households headed by women (top left), the average age of household heads (top right), and the proportion of heads that are literate (bottom left). We also provide the linear fit between each variable and the forcing variable on either side of the threshold. There are two features of each of the three graphs that apparently ensure that households do not systematically differ according to the municipio distance to threshold. First, the best fit lines almost coincide on either side of the threshold for each of the three variables. Second, in each case the “cloud” of municipio level data points appears continuous; there is no obvious structural break. In the bottom right corner, we include the logarithm of the municipio population on the right hand side, since we know that municipios entering in 2007 were generally larger in population than those entering in 2006. Again, the cloud of data appears to be contiguous around the threshold, indicating that there is not an important discontinuity at the threshold that might affect our results.

**Figure 6.1—Relationship between difference in distance to cluster centers and other variables measured at the *municipio* level in El Salvador**



Source: Authors' calculations based on El Salvador, Ministry of Economy 2007.

## 7. RESULTS

In this section, we first discuss results among primary-school-age children (7 to 12 years of age), then among preschool-age children (6-year-olds). We first use the data from the impact evaluation before describing results using the census data. We perform our main robustness checks using the census data because the amount of available data makes for richer tests.

### Impacts from the CSR Evaluation Data

Since the CSR evaluation data include measures of school enrollment in both 2006 and 2007, we construct an estimator of impacts using difference-in-differences. We use three different kernels to estimate equation (3): the rectangular kernel, the Gaussian kernel, and the Epanechnikov kernel. We also estimate the impacts using local linear regression; to do so, we allow complete flexibility in the slope of the relationship between school enrollment and the forcing variable:

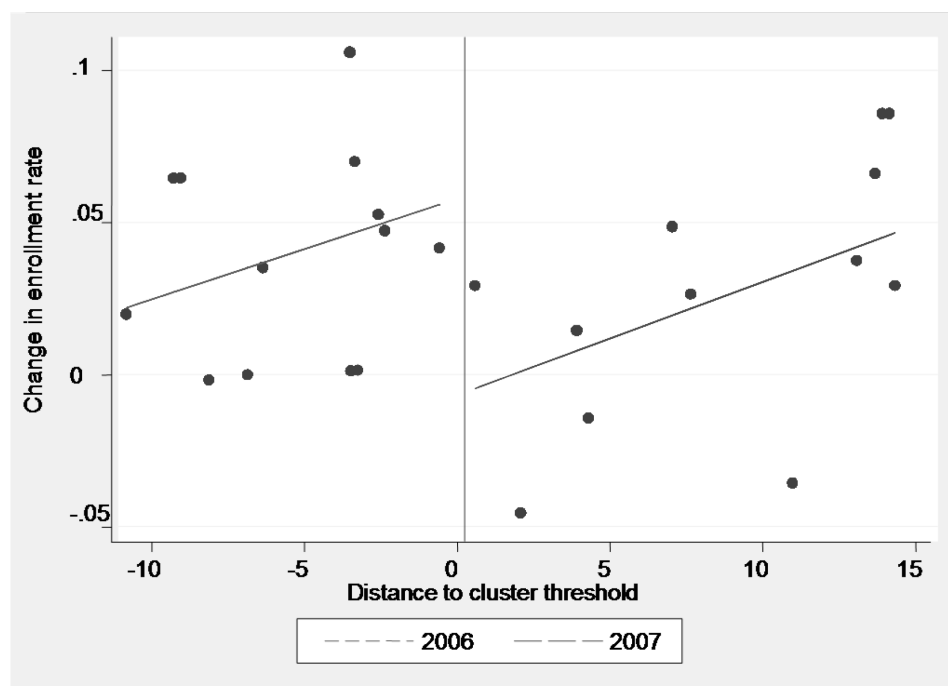
$$Y_i = \alpha + \beta_1 T_i + \beta_2 G_i + \beta_3 T_i G_i + \beta_4 D_i + \beta_5 T_i D_i + \beta_6 T_i G_i + \beta_7 T_i G_i + \varepsilon_i \forall |D_i| \leq h, \quad (9)$$

where  $T_i$  references time,  $G_i$  references the CSR entry group, and  $D_i$  represents the difference in distance between the cluster centers, which is defined as negative when closer to the 2006 entry group cluster center and positive when closer to the 2007 entry group cluster center. The variable  $h$  represents the bandwidth, which we vary in estimation. The estimate of impact is  $\beta_3$ , which estimates the difference in the slope after the program began (in 2007). The coefficients  $\beta_4$  through  $\beta_7$  represent the local slopes with respect to the difference in the distance to the two cluster centers and are allowed to vary over time and by entry group. When we estimate the impacts using the rectangular kernel, we use equation (9) and restrict  $\beta_4$  through  $\beta_7$  to zero.

We initially estimate the impacts of CSR on the enrollment of 7- to 12-year-olds in school based on the impact evaluation data. In doing so, we first difference *municipio* average enrollment rates and graph them against the difference in distance between cluster centers (Figure 7.1), along with a linear fit on both sides of the threshold. The figure shows a classic pattern of impact for regression discontinuity. On both sides of the implicit threshold, changes in enrollment rates are rising toward the threshold among CSR transfer recipients and then drop suddenly at the threshold and begin to rise again farther from the threshold.

When we estimate the impacts close to the threshold, we find reasonably strong evidence that CSR increased primary-school enrollment rates for 7- to 12-year-olds between 2006 and 2007 in *municipios* entering the program in 2006 (Table 7.1). Using the rectangular kernel, coefficient estimates increase as we narrow the bandwidth; the point estimate at the most narrow bandwidth suggests an increase in enrollment of 5.2 percentage points at the threshold, and it is significant at the 5 percent level. On the other hand, the coefficients estimated using local linear regression fall somewhat as the bandwidth narrows, though most are statistically significant at the 5 or 10 percent level. At the most narrow bandwidth, the estimate is very consistent with that of the rectangular kernel, at 4.7 percentage points. Using other nonparametric kernels, we also find similar results, so it is safe to conclude that the impact evaluation data suggest an impact of around 5 percentage points at the implicit threshold.

**Figure 7.1—Change in school enrollment rate in El Salvador from 2006 to 2007 among children 7 to 12 years of age by distance from implied cluster threshold, 2006 and 2007 entry groups**



Source: Authors' calculations based on Impact Evaluation Baseline Survey, *Comunidades Solidarias Rurales*, 2008.

**Table 7.1—Regression discontinuity results for impact of transfer associated with *Comunidades Solidarias Rurales* on change in school enrollment rates in El Salvador from 2006 to 2007 among children 7 to 12 years of age, comparing 2006 entrants with 2007 entrants**

	Full sample (1)	Euclidean distance bandwidth = 10 (2)	Euclidean distance bandwidth = 8 (3)	Euclidean distance bandwidth = 5 (4)
OLS (rectangular kernel)	0.015 (0.019)	0.031 (0.018)*	0.030 (0.020)	0.052 (0.023)**
Local linear regression	0.066 (0.028)**	0.071 (0.027)**	0.082 (0.026)***	0.047 (0.037)
<i>Nonparametric kernels</i>				
Gaussian	0.020 (0.016)	0.034 (0.020)*	0.035 (0.019)*	0.052 (0.021)**
Epanechnikov	0.030 (0.017)*	0.038 (0.020)*	0.045 (0.020)**	0.054 (0.021)***
Number of observations	3,239	2,306	2,105	1,534

Source: Impact Evaluation Baseline Survey, *Comunidades Solidarias Rurales*, 2008.

Notes: OLS: ordinary least squares. Standard errors in parentheses clustered at the *municipio* level. \* indicates significance at the 10 percent level; \*\* indicates significance at the 5 percent level; \*\*\* indicates significance at the 1 percent level. For kernel estimates, standard errors are bootstrapped using 100 replications of the data. Bandwidth refers to the distance on either side of the threshold; where a bandwidth is specified, any observations outside the bandwidth are excluded.

Given that enrollment rates were already above 90 percent on average prior to program entry, these estimates are relatively large. Whereas other similar conditional cash transfer (CCT) programs had slightly

larger impacts on enrollment among approximately the same age groups, these results are quite good among situations with similar baseline enrollment rates. To give specific examples, Maluccio and Flores (2005) estimated an impact of 12.8 percentage points among children 7 to 13 years of age in Nicaragua attributable to *Red de Protección Social*, but the baseline enrollment was 60.7 percent.<sup>12</sup> In Colombia, where baseline enrollment was 91.7 percent among 8- to 13-year-olds, *Familias en Acción* had an impact of 2.1 percentage points on enrollment (Attanasio, Fitzsimmons, and Gomez 2005), and no statistically significant impact was found among children in Mexico in primary-school grades due to PROGRESA (Schultz 2004). CSR has been very effective in getting the last few unenrolled children into primary school. In fact, findings from the 2006 entry group described here were mirrored in the 2007 and 2008 entry groups in follow-up surveys; by 2010, when the final evaluation survey was conducted, enrollment rates among 7- to 12-year-olds were close to 100 percent for all three represented entry groups (de Brauw et al. 2010).

### Impact Results Using the Population Census

To use the population census data to measure impacts, we can only generate single difference impact estimates. By comparing these estimates with estimates from the impact evaluation data, however, we can understand whether the results might be biased in either direction. Furthermore, the census data have two distinct advantages. First, the census data by definition include all *municipios*, so there are additional degrees of freedom we can add to the analysis. Second, the entire rural population of *municipios* entering in 2006 and 2007 is represented, so we can break up the population into groups by age and gender to understand who was experiencing the impact.

As with the impact estimates for the evaluation data, we use the same three kernels and local linear regression to estimate impacts. To estimate impacts using local linear regression, we modify equation (9) to fit the single difference available:

$$Y_i = \alpha + \beta_1 T_i + \beta_2 D_i + \beta_3 T_i D_i + \varepsilon_i \quad \forall |D_i| \leq h, \quad (10)$$

where  $Y$  again represents school enrollment,  $D_i$  is a dummy variable that represents the 2006 entry group, and  $D_i$  represents the difference in distance between cluster centers. In some specifications, we also include a vector of child characteristics,  $X_i$ , which includes age and gender. The coefficient  $\beta_1$  now represents the impact of the transfer associated with *Comunidades Solidarias Rurales*, and we use both the rectangular kernel and local linear regression to estimate equation (10). The rectangular kernel implies that  $\beta_2 = \beta_3 = 0$ .

We initially graph average enrollment rates among 7- to 12-year-olds at the *municipio* level for 2007 (Figure 7.2). The figure clearly suggests impact. The enrollment rates in 2006 are almost all between 90 percent and 100 percent, and with the exception of one outlier, they are tightly distributed. On the right-hand side of the threshold (2007 entry group), enrollment rates are much more variable and are clearly lower on average. One outlier is clearly much lower than the other *municipios* as well. In general, however, this graph is consistent with the evaluation data, in that it is highly suggestive of impacts on the net enrollment rate.

Next, we estimate equation (10) using the rectangular kernel, local linear regression, and the two nonparametric kernels (Table 7.2). The initial estimate implies that among the entire population of children in the 2006 entry group, school enrollment is 6.9 percentage points higher than among children in the 2007 entry group. However, this estimate includes a number of *municipios* that are not very close to the threshold; when we use local linear regression to account for any linear effects the distance from the threshold might have on this impact estimate, the estimate drops to 4.6 percent. Perhaps our best estimate of the impact at the threshold, however, comes at the most narrow bandwidth; the results suggest an impact of

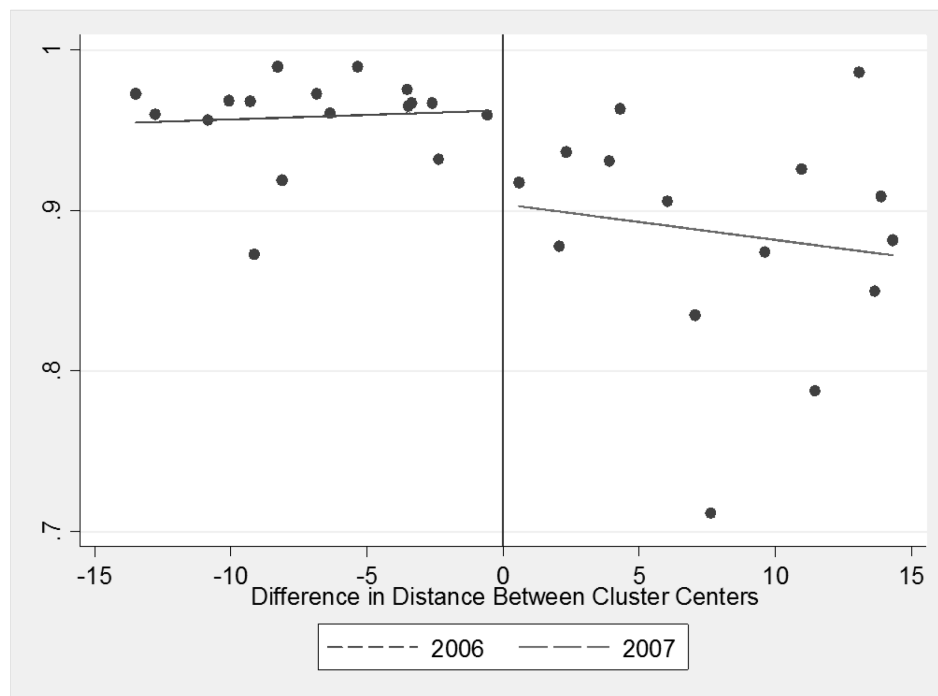
---

<sup>12</sup>Similar impacts were found in Chile among 6- to 15-year-olds due to *Chile Solidario* (7.5 percent; Galasso 2006) and in Ecuador due to *Bono de Desarrollo Humano* (10.3 percent; Schady and Araujo 2008).



between 3.7 and 4.0 percentage points, depending upon the estimation strategy. Only the local linear regression results are not significant at the 5 percent level or better.

**Figure 7.2—Average net school enrollment rates in El Salvador in 2007 at the *municipio* level among children 7 to 12 years of age, comparing 2006 entry group with 2007 entry group**



Source: Authors' calculations based on El Salvador, Ministry of Economy 2007.

**Table 7.2—Regression discontinuity results for impact of transfer associated with *Comunidades Solidarias Rurales* on school enrollment rates in El Salvador in 2007 among children 7 to 12 years of age, comparing 2006 entrants with 2007 entrants**

	Full sample (1)	Euclidean distance bandwidth = 10 (2)	Euclidean distance bandwidth = 5 (3)
OLS (rectangular kernel)	0.069 (0.016)**	0.059 (0.018)**	0.037 (0.012)**
Local linear regression	0.047 (0.017)**	0.030 (0.019)	0.038 (0.021)*
<i>Nonparametric kernels</i>			
Gaussian	0.068 (0.015)**	0.055 (0.019)**	0.038 (0.015)**
Epanechnikov	0.067 (0.015)**	0.050 (0.017)**	0.040 (0.016)**
Number of observations	37,221	21,777	14,528

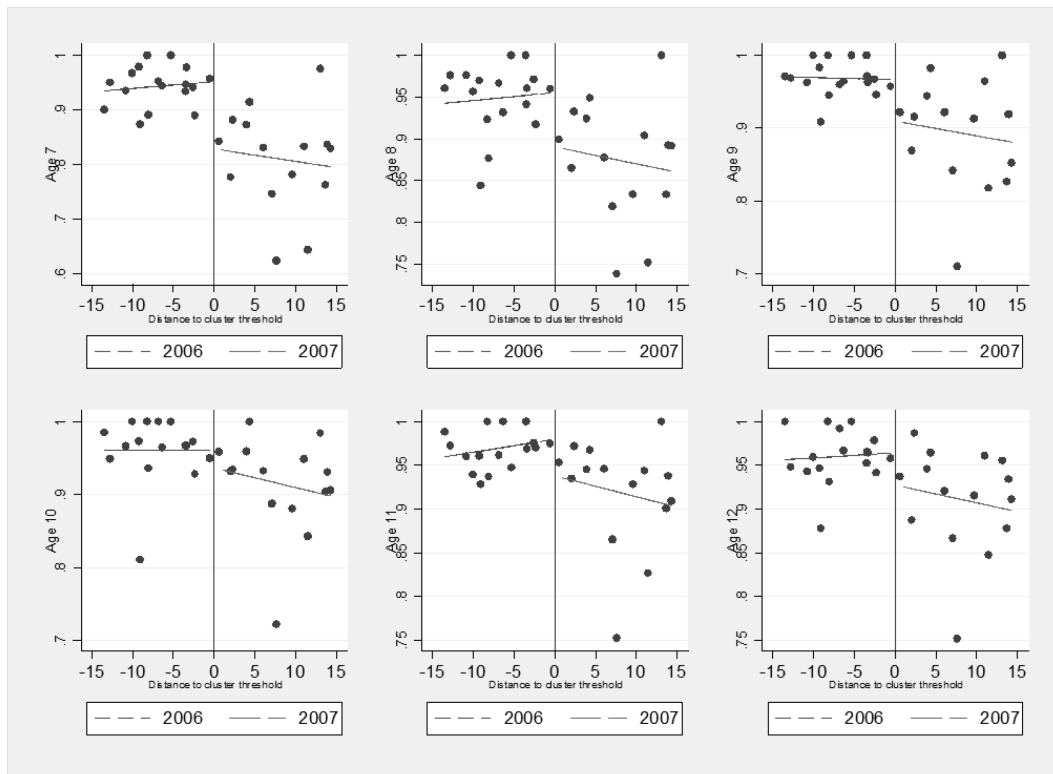
Source: El Salvador, Ministry of Economy 2007.

Notes: OLS: ordinary least squares. Standard errors clustered at the *municipio* level are in parentheses. \* indicates significance at the 10 percent level; \*\* indicates significance at the 5 percent level. For kernel estimates, standard errors are bootstrapped using 100 replications of the data. All regressions include a full set of age and gender dummies.

We can conservatively conclude, then, that the local average treatment effect of the transfer associated with CSR at the threshold is 3.7 percentage points. Yet there is enough data in the census, by definition, to potentially better isolate whether increased enrollment is taking place among older or younger children, or among boys or girls. Therefore we next investigate graphically and with regressions the ages among which increased enrollment occurred, and whether it occurred among boys, girls, or both.

We initially plot average enrollment rates, by age and by the difference in distance between cluster centers (Figure 7.3). The figure suggests that impacts have occurred among younger age groups but not as clearly among older age groups. For example, all enrollment rates for 7-year-olds in the 2006 entry group are again tightly distributed between 90 percent and 100 percent, whereas they are quite dispersed and much lower on average for 7-year-olds in the 2007 entry group. Yet for 10-year-olds, for example, there is less dispersion in the 2007 entry group, although there are a few outliers. Almost all of the *municipios*, however, appear to have enrollment rates that are between 90 percent and 100 percent. Therefore we expect to observe larger impacts on younger children than on older children.

**Figure 7.3—Average net enrollment rates in El Salvador among children 7 to 12 years of age, by age and *municipio* level, comparing 2006 entry group with 2007 entry group**



Source: Authors' calculations based on El Salvador, Ministry of Economy 2007.

We next estimate the impacts of the transfer associated with CSR on school enrollment rates by age and gender, using only the narrowest bandwidth and using both the rectangular kernel and local linear regression (Table 7.3). The results are consistent with the illustrations above but also reveal some interesting gender differences. First, we find that the impact is largest among 7-year-olds. According to the rectangular kernel results, among 7-year-olds enrollment is 8.9 percentage points higher in the 2006 entry group than in the 2007 entry group due to the transfer payment. The coefficient estimates for all children remain significant but decline in magnitude until age 9. At age 10, the coefficient estimates are insignificant among both estimation methods, and among 11- and 12-year-olds, they are both only significant for the

rectangular kernel, and both are reasonably small in magnitude (around 2.5 percentage points for both). Therefore it is clear that the largest impacts are among younger children.

**Table 7.3—Impact of *Comunidades Solidarias Rurales* on school enrollment in El Salvador, by age and gender, bandwidth of 5**

Age	All children		Boys		Girls	
	OLS	LLR	OLS	LLR	OLS	LLR
7	0.089 (0.023)**	0.098 (0.046)**	0.096 (0.026)**	0.117 (0.035)**	0.081 (0.025)**	0.074 (0.067)
8	0.042 (0.015)**	0.054 (0.024)**	0.037 (0.016)**	0.051 (0.026)*	0.047 (0.016)**	0.058 (0.024)**
9	0.039 (0.013)**	0.04 (0.019)**	0.033 (0.010)**	0.022 (0.014)	0.047 (0.017)**	0.06 (0.026)**
10	-0.001 (0.012)	-0.019 (0.021)	0.001 (0.021)	0.019 (0.028)	-0.004 (0.017)	-0.048 (0.023)*
11	0.026 (0.006)**	0.015 (0.011)	0.018 (0.009)*	-0.03 (0.007)	0.033 (0.013)**	0.067 (0.019)**
12	0.024 (0.014)*	0.023 (0.017)	0.014 (0.011)	<0.001 (0.017)	0.034 (0.019)	0.047 (0.021)**

Source: El Salvador, Ministry of Economy 2007.

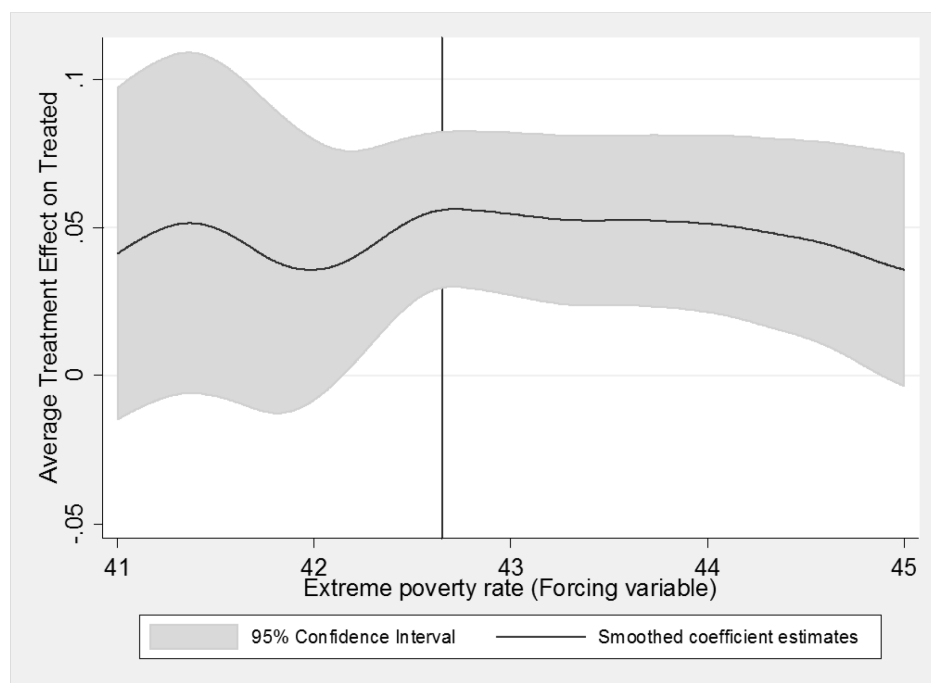
Notes: OLS: ordinary least squares. LLR: local linear regression. Standard errors clustered at *municipio* level in parentheses. Each cell represents a separate regression. \* indicates significance at the 10 percent level; \*\* indicates significance at the 5 percent level. For kernel estimates, standard errors are bootstrapped using 100 replications of the data. Regressions compare individuals in 2006 entry *municipios* with those in 2007 entry *municipios*.

These results have two main implications. First, they suggest that a major impact of the CSR transfers was children's enrolling in school earlier than they might have otherwise; in relative terms, the impacts on enrollment among 7-year-olds were quite large. Second, results suggest that older girls became slightly more likely either to stay in school or to initially enter school as a consequence of the household's receiving a transfer. The former point is important because it foreshadows lower repeat rates, which have been correlated with receiving the transfer and appear to be correlated with earlier school entry, which is a direct consequence of the program. The latter point is important because the impact of CSR on older girls is clearly different from the impact on older boys.

### Robustness Check: Alternative Forcing Variable

We also estimate the impacts on school enrollment using an alternative forcing variable, the poverty rate used in determining the cluster to which each *municipio* belonged, again using the 2007 census data. As with the implicit partition, we test two different bandwidths, a relatively large bandwidth (8) and a narrower one (3), along with estimating impacts using the full sample. We find the midpoint between the lowest poverty rate in the treatment group (43.1) and the highest poverty rate in the control group (42.2) and use that rate, 42.65, as an initial threshold. In the Appendix (Table A.1), we present estimates using the three kernels and local linear regression; in particular, it is worth highlighting that using any of the three kernel estimates and the most narrow bandwidth, we come up with an impact estimate of 3.7 percentage points, exactly as we did with the implicit partition. We also use the rectangular kernel and the narrow bandwidth to search over all possible thresholds between poverty rates of 41 and 45 (Figure 7.4). Between about 42.2 and 43.1, all impact estimates are statistically different from zero at the 5 percent level or better. The results from using the alternative forcing variable are therefore quite consistent with results from the difference in distance between cluster centers.

**Figure 7.4—Smoothed impact estimates using alternative forcing variable (poverty rate), children 7 to 12 years of age in El Salvador**



Source: Authors calculations based on El Salvador, Ministry of Economy 2007.

Notes: Vertical line indicates the midpoint between the lowest poverty rate in the 2006 entry group and the highest poverty rate in the 2007 entry group.

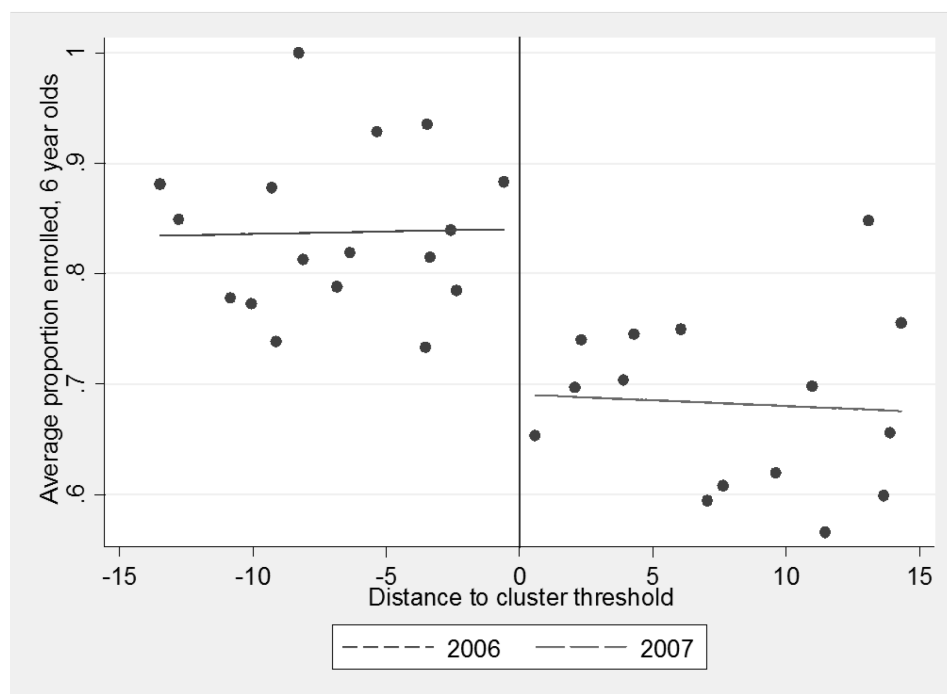
### Impact Estimates on *Parvularia* Attendance

As discussed in the descriptive section, enrollment rates among 6-year-olds, who are of age to attend *parvularia*, were much higher among children in the 2006 entry group than among those in the 2007 entry group. Using the census data, we find that the difference in enrollment rates is also apparent when we graph *municipio* averages by the difference in distance between cluster centers (Figure 7.5). School enrollment rates among 6-year-olds were fairly consistently high in the 2006 entry group, with most above 80 percent. Among the 2007 entry group, the rates were more variable, but almost all were below 80 percent, with one below 60 percent. Graphically, there is fairly clear evidence of an impact on school enrollment among 6-year-olds.

To more precisely estimate the impacts of CSR on school enrollment among 6-year-olds, we follow the same estimation strategy as for primary-school enrollment, using the 2007 census data (Table 7.4). Not surprisingly, we find large impact estimates when we simply estimate equation (10) among 6-year-olds without restricting the sample; the estimates are 16.9 percentage points at the mean and 15.3 percentage points using local linear regression (column 1). When we restrict the bandwidth to 5, the estimate using the rectangular kernel drops slightly to 14.8 percentage points, but the local linear regression result increases to 19.6 percentage points (column 3). The effect seems to be larger among girls (column 5) than among boys (column 4): Examining the results from the three kernels shows an impact of around 13 percentage points among boys and between 16 and 17 percentage points among girls. Results for the local linear regression are even higher at the narrow threshold, although also more imprecise. These results imply that even if CSR has not had large impacts on primary-school enrollment—in part because large impacts were not possible given that enrollment rates prior to CSR were quite high—still, a larger proportion of children in poorer

*municipios* will have some preschool experience, which may imply improved grade progression in the future.

**Figure 7.5.—Average net enrollment rates, 6 year olds, Municipio level, Comparing 2006 to 2007 Entry group, El Salvador**



Source: Authors' calculations based on El Salvador, Ministry of Economy 2007.

**Table 7.4—Regression discontinuity results for impact of transfer associated with *Comunidades Solidarias Rurales* on school enrollment rates in El Salvador in 2007 among children 6 years of age, comparing 2006 entrants with 2007 entrants**

	Full sample	Bandwidth = 10		Bandwidth = 5	
	All children (1)	All children (2)	All children (3)	Boys (4)	Girls (5)
Rectangular kernel	0.169 (0.019)**	0.159 (0.021)**	0.148 (0.029)**	0.132 (0.040)**	0.164 (0.033)**
Local linear regression	0.153 (0.032)**	0.149 (0.041)**	0.196 (0.060)**	0.161 (0.080)*	0.238 (0.046)**
<i>Nonparametric kernels</i>					
Gaussian	0.168 (0.021)**	0.158 (0.024)**	0.151 (0.030)**	0.134 (0.038)**	0.169 (0.034)**
Epanechnikov	0.166 (0.021)**	0.153 (0.026)**	0.155 (0.031)**	0.135 (0.039)**	0.176 (0.037)**
Number of observations	6,209	3,665	2,509	1,294	1,209

Source: El Salvador, Ministry of Economy 2007.

Notes: Standard errors clustered at the *municipio* level are in parentheses. \* indicates significance at the 10 percent level; \*\* indicates significance at the 5 percent level. For kernel estimates, standard errors are bootstrapped using 100 replications of the data.

## Final Robustness Check: Adding Household Characteristics

As a final robustness check, we re-estimate local linear regressions using the narrow bandwidth for both 6-year-olds and 7- to 12-year-olds, including additional child-level and household-level characteristics (Table 7.5). We expect that even if some of the additional characteristics are significantly related to school enrollment, they should be orthogonal to the impact at the threshold and therefore should not affect impact estimates. Our results are quite consistent with expectations. Among both 6-year-olds and 7- to 12-year-olds, having a literate head of household was associated with higher probability of school enrollment; however, the inclusion of the additional variables does not affect impact estimates (columns 3 and 6). In fact, we find that among 7- to 12-year-olds, the additional variables explain enough variation in the data to make the coefficient estimate significant at the 5 percent rather than the 10 percent level, without changing the point estimate. In summary, the results are clearly robust to the inclusion of individual or household characteristics.

**Table 7.5—Regression discontinuity results for impact of transfer associated with Comunidades Solidarias Rurales on school enrollment rates in El Salvador in 2007 among children 6 years of age and children 7 to 12 years of age, controlling for individual and household characteristics, bandwidth of 5**

	Age 6			Ages 7–12		
	(1)	(2)	(3)	(4)	(5)	(6)
Impact at threshold	0.196 (0.061)**	0.196 (0.061)**	0.191 (0.063)**	0.038 (0.021)*	0.036 (0.021)	0.038 (0.017)**
Difference in distance between cluster centers	0.020 (0.005)**	0.020 (0.005)**	0.020 (0.005)**	0.006 (0.004)	0.006 (0.004)	0.005 (0.003)
Difference in distance * threshold	-0.016 (0.025)	-0.015 (0.025)	-0.017 (0.026)	-0.011 (0.007)	-0.010 (0.007)	-0.010 (0.005)*
Gender (1 = male)		-0.006 (0.022)	-0.009 (0.020)		0.007 (0.005)	0.007 (0.005)
Age (in years)					0.014 (0.003)**	0.014 (0.003)**
Female-headed household (1 = yes)			0.021 (0.019)			0.010 (0.005)*
Head is literate (1 = yes)			0.065 (0.012)**			0.054 (0.009)**
Age of household head			-0.002 (0.008)			0.004 (0.001)**
Number of observations	2,529	2,529	2,517	14,528	14,528	14,467

Source: El Salvador, Ministry of Economy 2007.

Notes: Standard errors clustered at the *municipio* level are in parentheses. \* indicates significance at the 10 percent level; \*\* indicates significance at the 5 percent level.

## 8. CONCLUSION

In this paper, we first developed conditions necessary to locally estimate program impacts using regression discontinuity when an explicit forcing variable is not used to determine program eligibility. In the case of CSR, partitioned cluster analysis was used to assign *municipios* to specific clusters. We show that under a reasonable set of assumptions, an implicit threshold exists between clusters and the distance to that threshold can be used as a forcing variable. We recommend using the same distance measure that is used in partitioned cluster analysis as the measure of the forcing variable.

We then apply this strategy to estimate the impacts on school enrollment among primary-school-aged children (7- to 12-year-olds) in rural El Salvador. We use two different datasets, one specifically collected for the impact evaluation, the other from El Salvador's 2007 national census, to show that impacts on primary-school enrollment were between 3.7 and 5.2 percentage points, depending upon the sample and estimation strategy. These impacts are relatively large given that enrollment was already above 90 percent for this age range. Our impact estimates are robust to using a more traditional forcing variable, the poverty rate, because the clustering appears to have occurred mainly along the poverty rate rather than the severe stunting rate, likely due to the fact that the poverty rate varied more than the severe stunting rate. Results are also robust to the inclusion of individual or household characteristics in a local linear regression framework. Using the same estimation framework, we find even larger impact estimates among 6-year-olds, who are of age to attend *parvularia*. We attribute an increase of about 15 percentage points in school enrollment among 6-year-olds to the transfers associated with CSR.

From a policy perspective, one might argue that it would have been a better use of resources to target middle-school enrollment rather than primary-school enrollment. According to school censuses collected between 2005 and 2009, in *municipios* with households receiving transfers from CSR, 9th grade enrollment was 36 percent lower than 6th grade enrollment had been three years previously, indicating that a large proportion of children dropped out of school between 6th grade and 9th grade. Middle-school targeting could have had larger overall impacts on school enrollment. However, it may be particularly important that school enrollment among younger children has substantially increased as a consequence of CSR. El Salvador has long had very high first-grade repetition, on the order of 22 percent nationwide in the data collected between 2001 and 2004 to generate a poverty map. To the extent that earlier enrollment is negatively correlated with first-grade repetition (as argued in IFPRI and FUSADES 2009), one might expect repetition rates to decline as children enter school earlier in El Salvador. The simplicity of targeting in CSR has made it reasonably easy to implement, and more complicated categorical targeting might have made the program more costly to implement well. For example, targeting both *parvularia* enrollment among 6-year-olds and middle-school enrollment might have confused implementing officials about the eligibility of certain children, which is easy to determine in the program's current form.

## APPENDIX: SUPPLEMENTARY TABLE

**Table A.1—Regression discontinuity results for impact of transfer associated with *Comunidades Solidarias Rurales* on school enrollment rates in El Salvador in 2007 among children 7 to 12 years of age, comparing 2006 entrants with 2007 entrants, using poverty rate as the forcing variable**

	Full sample (1)	Euclidean distance bandwidth = 10 (2)	Euclidean distance bandwidth = 5 (3)
OLS (rectangular kernel)	0.069 (0.016)**	0.070 (0.016)**	0.037 (0.012)**
Local linear regression	0.054 (0.019)**	0.052 (0.019)**	0.031 (0.015)*
<i>Nonparametric kernels</i>			
Gaussian	0.069 (0.017)**	0.068 (0.018)**	0.037 (0.015)**
Epanechnikov	0.069 (0.017)**	0.066 (0.019)**	0.037 (0.015)**
Number of observations	37,221	34,865	14,528

Source: Authors' calculations based on El Salvador, Ministry of Economy 2007.

Notes: OLS: ordinary least squares. Standard errors clustered at the *municipio* level are in parentheses. \* indicates significance at the 10 percent level; \*\* indicates significance at the 5 percent level. For kernel estimates, standard errors are bootstrapped using 100 replications of the data. All regressions include a full set of age and gender dummies.



## REFERENCES

- Attanasio, O., E. Fitzsimmons, and A. Gomez. 2005. *The Impact of a Conditional Education Subsidy on School Enrollment in Colombia*. Report Summary: *Familias* 01. London: Institute for Fiscal Studies.
- Battistin, E., A. Brugiavani, E. Rettore, and G. Weber. 2009. "The Retirement Consumption Puzzle: Evidence from a Regression Discontinuity Approach." *American Economic Review* 99 (5): 2209–2226.
- Bayer, P., F. Ferreira, and R. McMillan. 2007. "A Unified Framework for Measuring Preferences for Schools and Neighborhoods." *Journal of Political Economy* 115 (4): 588–638.
- Behrman, J., and E. King. 2009. "Timing and Duration of Exposure in Evaluations of Social Programs." *World Bank Research Observer* 24 (1): 55–82.
- Black, S. 1999. "Do Better Schools Matter? Parental Valuation of Elementary Education." *Quarterly Journal of Economics* 114 (2): 577–599.
- Buddelmeyer, H., and E. Skoufias. 2003. *An Evaluation of the Performance of Regression Discontinuity Design on PROGRESA*. Discussion Paper 827. Bonn, Germany: Institute for the Study of Labor (IZA).
- Calinski, R.B., and J. Harabasz, 1974, "A Dendrite Method for Cluster Analysis," *Communications in Statistics* 3: 1-27.
- Chaudhury, N., and D. Parajuli. 2006. "Conditional Cash Transfers and Female Schooling: The Impact of the Female School Stipend Program on Public School Enrollments in Punjab, Pakistan." World Bank Policy Research Working Paper no. 4102, Washington, DC.
- de Brauw, A., M. Adato, A. Peterman, T. Roopnaraine, M. B. de Sanfeliu, M. Shi, R. Pleitez, et al. 2010. "Informe de Sostenibilidad del Programa." Report submitted to FISDL (*Fondo de Inversión Social para el Desarrollo Local*). Mimeo, San Salvador, El Salvador.
- Edmonds, E., K. Mammen, and D. Miller. 2005. "Rearranging the Family? Income Support and Elderly Living Arrangements in a Low Income Country." *Journal of Human Resources* 40 (1): 186–207.
- El Salvador, Fondo de Inversión Social para el Desarrollo Local (FISDL), 2004. *Mapa de Pobreza*, San Salvador, El Salvador.
- El Salvador, Ministry of Economy. 2007. *Censo Nacional de Población y Vivienda de 2007*. San Salvador.
- Filmer, D., and N. Schady. 2009. "Who Benefits? Scholarships, School Enrollment and Work of Recipients and Their Siblings." Unpublished manuscript, World Bank, Washington, DC.
- Galasso, E. 2006. "With Their Effort and One Opportunity: Alleviating Extreme Poverty in Chile." Unpublished manuscript, Development Research Group, World Bank, Washington, DC.
- IFPRI (International Food Policy Research Institute) and FUSADES (*Fundación Salvadoreña para el Desarrollo Económico y Social*). 2009. "Informe de Impactos al Año de Implementación." Report submitted to FISDL (*Fondo de Inversión Social para el Desarrollo Local*). Mimeo, San Salvador, El Salvador.
- Imbens, G., and T. Lemieux. 2008. "Regression Discontinuity Designs: A Guide to Practice." *Journal of Econometrics* 142 (2): 807–828.
- Imbens, G., and J. Wooldridge. 2009. "Recent Developments in the Econometrics of Program Evaluation." *Journal of Economic Literature* 47 (1): 5–86.
- "Impact Evaluation Baseline Survey, *Comunidades Solidarias Rurales*." 2008. San Salvador, El Salvador: IFPRI (International Food Policy Research Institute) and FUSADES (*Fundación Salvadoreña para el Desarrollo Económico y Social*).
- Kane, T. 2003. *A Quasi-Experimental Estimate of the Impact of Financial Aid on College Going*. Working Paper no. 9703. Cambridge, Mass.: National Bureau of Economic Research.
- Lalive, Rafael, 2008, "How do Extended Benefits Affect Unemployment Duration? A Regression Discontinuity Approach," *Journal of Econometrics* 142(2): 785-806.

- Lavy, V. 2006. *From Forced Busing to Free Choice in Public Schools: Quasi-Experimental Evidence of Individual and General Effects*. Working Paper no. 11969. Cambridge, Mass.: National Bureau of Economic Research.
- Lee, D., and T. Lemieux. 2010. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature* 48 (2): 281–355.
- Ludwig, J., and D. Miller. 2007. "Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design." *The Quarterly Journal of Economics* 122 (1): 159–208.
- Maluccio, J., and R. Flores. 2005. *Impact Evaluation of a Conditional Cash Transfer Program: The Nicaraguan Red de Protección Social*. Research Report 141. Washington, DC: International Food Policy Research Institute.
- Papay, John P., John B. Willett, and Richard J. Murnane, 2011, "Extending the Regression-Discontinuity Approach to Multiple Assignment Variables," *Journal of Econometrics* 161(2): 203-207.
- Pence, K., 2006, "Foreclosing on Opportunity: State Laws and Mortgage Credit," *Review of Economics and Statistics* 88(February): 177-182.
- Ponce, J., and A. Bedi. 2010. "The Impact of a Cash Transfer Program on Cognitive Achievement: The *Bono de Desarrollo Humano* of Ecuador." *Economics of Education Review* 29:116–125.
- Porter, J. 2003. "Estimation in the Regression Discontinuity Model." Working paper, University of Wisconsin, Madison.
- Schady, N., and M. C. Araujo. 2008. "Cash Transfers, Conditions, and School Enrollment in Ecuador." *Economia* 8 (2): 43–70.
- Schultz, T. P. 2004. "School Subsidies for the Poor: Evaluating the Mexican PROGRESA Poverty Program." *Journal of Development Economics* 74 (1): 199–250.

## RECENT IFPRI DISCUSSION PAPERS

For earlier discussion papers, please go to <http://www.ifpri.org/publications/results/taxonomy%3A468>. All discussion papers can be downloaded free of charge.

1115. *The quiet revolution in India's food supply chains*. Thomas Reardon and Bart Minten, 2011.
1114. *A review of input and output policies for cereals production in Nepal*. Hemant Pullabhotla, Ganga Shreedhar, A. Ganesh-Kumar, and Ashok Gulati, 2011.
1113. *Do shocks affect men's and women's assets differently?: A review of literature and new evidence from Bangladesh and Uganda*. Agnes R. Quisumbing, Neha Kumar, and Julia A. Behrman, 2011.
1112. *Overcoming successive bottlenecks: The evolution of a potato cluster in China*. Xiaobo Zhang and Dinghuan Hu, 2011.
1111. *The impact of land titling on labor allocation: Evidence from rural Peru*. Eduardo Nakasone, 2011.
1110. *A multiregion general equilibrium analysis of fiscal consolidation in South Africa*. Margaret Chitiga, Ismael Fofana, and Ramos Mabugu, 2011.
1109. *How far do shocks move across borders?: examining volatility transmission in major agricultural futures markets*. Manuel A. Hernandez, Raul Ibarra, and Danilo R. Trupkin, 2011.
1108. *Prenatal seasonality, child growth, and schooling investments: Evidence from rural Indonesia*. Futoshi Yamauchi, 2011.
1107. *Collective Reputation, Social Norms, and Participation*. Alexander Saak, 2011.
1106. *Food security without food transfers?: A CGE analysis for Ethiopia of the different food security impacts of fertilizer subsidies and locally sourced food transfers*. A. Stefano Caria, Seneshaw Tamru, and Gera Bizuneh, 2011.
1105. *How do programs work to improve child nutrition?: Program impact pathways of three nongovernmental organization intervention projects in the Peruvian highlands*. Sunny S. Kim, Jean-Pierre Habicht, Purnima Menon, and Rebecca J. Stoltzfus, 2011.
1104. *Do marketing margins change with food scares?: Examining the effects of food recalls and disease outbreaks in the US red meat industry*. Manuel Hernandez, Sergio Colin-Castillo, and Oral Capps Jr., 2011.
1103. *The seed and agricultural biotechnology industries in India: An analysis of industry structure, competition, and policy options*. David J. Spielman, Deepthi Kolady, Anthony Cavalieri, and N. Chandrasekhara Rao, 2011.
1102. *The price and trade effects of strict information requirements for genetically modified commodities under the Cartagena Protocol on Biosafety*. Antoine Bouët, Guillaume Gruère, and Laetitia Leroy, 2011.
1101. *Beyond fatalism: An empirical exploration of self-efficacy and aspirations failure in Ethiopia*. Tanguy Bernard, Stefan Dercon, and Alemayehu Seyoum Taffesse, 2011.
1100. *Potential collusion and trust: Evidence from a field experiment in Vietnam*. Maximo Torero and Angelino Viceisza, 2011.
1099. *Trading in turbulent times: Smallholder maize marketing in the Southern Highlands, Tanzania*. Bjorn Van Campenhout, Els Lecoutere, and Ben D'Exelle, 2011.
1098. *Agricultural management for climate change adaptation, greenhouse gas mitigation, and agricultural productivity: Insights from Kenya*. Elizabeth Bryan, Claudia Ringler, Barrack Okoba, Jawoo Koo, Mario Herrero, and Silvia Silvestri, 2011.
1097. *Estimating yield of food crops grown by smallholder farmers: A review in the Uganda context*. Anneke Fermont and Todd Benson, 2011.
1096. *Do men and women accumulate assets in different ways?: Evidence from rural Bangladesh*. Agnes R. Quisumbing, 2011.
1095. *Simulating the impact of climate change and adaptation strategies on farm productivity and income: A bioeconomic analysis*. Ismaël Fofana, 2011.
1094. *Agricultural extension services and gender equality: An institutional analysis of four districts in Ethiopia*. Marc J. Cohen and Mamusha Lemma, 2011.
1093. *Gendered impacts of the 2007–08 food price crisis: Evidence using panel data from rural Ethiopia*. Neha Kumar and Agnes R. Quisumbing, 2011.

**INTERNATIONAL FOOD POLICY  
RESEARCH INSTITUTE**

**[www.ifpri.org](http://www.ifpri.org)**

**IFPRI HEADQUARTERS**

2033 K Street, NW  
Washington, DC 20006-1002 USA  
Tel.: +1-202-862-5600  
Fax: +1-202-467-4439  
Email: [ifpri@cgiar.org](mailto:ifpri@cgiar.org)