HEDG

HEALTH, ECONOMETRICS AND DATA GROUP

THE UNIVERSITY of York

WP 11/16

The Effect of Medicaid on Children's Health: A Regression Discontinuity Approach

Dolores de la Mata

July 2011

york.ac.uk/res/herc/hedgwp

The Effect of Medicaid on Children's Health: A Regression Discontinuity Approach

Dolores de la Mata[*][†]

This version: July 2011

Abstract

In this paper I estimate the impact of Medicaid on children's health care utilization and their subsequent health outcomes. I estimate the causal effects using a Regression Discontinuity (RD) design. I exploit the discontinuity generated by Medicaid's eligibility rule, based on family income, on program participation rates. In contrast with a standard regression discontinuity approach, here there are multiple eligibility thresholds that vary across states. This feature allows me to estimate heterogeneous effects of the program at different income thresholds. Using data from the Panel Study of Income Dynamics (PSID) and its Child Development Study (CDS) supplement, I find that the effects of Medicaid on measures of children's health are heterogeneous depending on the family income level. Negative impacts of Medicaid are generally observed for children of higher-income families −between 185% and 250% of the poverty line−, while generally null or positive effects are observed for poorer children −family income between 100% and 185% of the poverty line. A possible explanation for the heterogeneous impacts is the differential effect of Medicaid on preventive health care utilization. While I find that Medicaid increases the use of preventive medical care among children with low family income, no significant effects are observed among those with higher income. Another likely explanation for the observed effects is that Medicaid induces higher-income families to drop private health insurance with access to better quality of health care, generating a negative effect on children's health outcomes.

JEL Classification: I18, G22.

Keywords: Health insurance, children's health, Medicaid, regression discontinuity design.

1 Introduction

There is strong evidence showing a positive relationship between parental socioeconomic status and children's health, leading to health inequalities in early childhood. To the extent that poor

[*]Department of Economics, Universidad Carlos III de Madrid. Email: mlmata@eco.uc3m.es.

[†]I am grateful to Matilde Machado for her constant support, encouragement and valuable advice. I also want to thank Lucila Berniell, Julio Cáceres, Joaquín Coleff, Ramiro de Elejalde, Roger Feldman, Eva García Morán, Daniel García, Felix Lobo, and seminar participants at the UC3M Student Workshop, UC3M Applied Economics and Econometrics seminar, Sao Paulo University, Universidad de Santiago de Chile, Universitat Rovira i Virgili, Universitat de les Illes Balears, the Office of Evaluation of the Inter-American Development Bank, and the 2nd IRDES Workshop of Applied Health Economics and Policy Evaluation for helpful comments and discussions. All remaining errors are my own.

health affects the formation of human capital, health may play a key role in the intergenerational transmission of socioeconomic inequalities (Currie, 2009; Almond and Currie, 2010). Currie (2009) suggests that children's health inequalities may be partially explained by disparities in the access to health care services. The provision of public health insurance coverage to children in low income families facilitates the access to medical care and, therefore, may help to weaken the link between socioeconomic status and health.

The US does not have a universal health care system which makes family income an important factor determining access to health care. Public health insurance programs in the US are designed to improve the access to medical care for low income individuals. Medicaid is a means-tested program and entitles those meeting the required conditions to have public health insurance coverage (Kaiser Commission on Medicaid and the Uninsured, 2010). Medicaid is the largest source of insurance coverage for children in the US, covering about 30% of all children and 59% of low income children.[1]

In this paper, I address three questions. First, I study whether Medicaid contributes to enhance children's utilization of health care services, and, more important, whether it contributes to improve their health outcomes. Second, I analyze whether Medicaid has lagged effects over health. Since health is a stock, the effects of insurance coverage may not be visible immediately but with some lag. Finally, I investigate whether the provision of free health insurance to relatively high income families can have some unintended negative effect on their children's health. The provision of a public free health insurance to children in certain ranges of family income may compete with private insurance and induce some of these families to drop the private alternative. If switching from the private to the public occurs then this could have a negative effect on children's health, as long as the switch implies a reduction in the quality of health care.

I exploit the particular characteristics of Medicaid's eligibility rules to identify the causal effects of the program on children's outcomes. A child is eligible to receive Medicaid coverage if his family income, as a percentage of the federal poverty line, is below a given threshold. This rule generates a discontinuity in the enrollment rates of children with family income close to the threshold, which allows me to implement a regression discontinuity (RD) design.

The eligibility criteria for the Medicaid program are set at the state level, therefore the income threshold that determines eligibility varies among states and has been changing through time. With multiple thresholds, the effects estimated pooling all thresholds are not restricted to the individuals located around a single income threshold, but they are averages of the effects across the different thresholds (Black et al., 2005; Bloom, 2009; Carneiro and Ginja, 2009). The multiplicity of thresholds also allows me to investigate whether the effects of Medicaid are heterogeneous for the different targeted levels of family income.

I use data from the Panel Study of Income Dynamics (PSID) and the Child Development Study (CDS) supplement, which provide rich information about children's health and health care utilization as well as detailed information on socioeconomic characteristics of the family. The PSID data allows tracking of children's Medicaid status at different ages through childhood.

[1] Low income children are those with family income below 200% of the federal poverty line. Source: Urban Institute and Kaiser Commission on Medicaid and the Uninsured estimates based on the Census Bureau's March 2009 and 2010 Current Population Survey (CPS: Annual Social and Economic Supplements). `http://www.statehealthfacts.org`.

In the first part of the paper (Section 5) I test the internal validity of the RD design by performing a number of checks which support the RD local assumption stating that eligibility is randomly assigned in the neighborhood of the thresholds. First, I show there is no evidence that families have perfect control over their income so that their children just qualify for Medicaid. Second, I show that the eligibility rule generates a discontinuity in Medicaid enrollment rates at the threshold. Moreover, my results indicate that this discontinuity is higher the lower the family income threshold. Third, I provide evidence that the discontinuity in participation rates at the threshold is not generated by discontinuous changes of other individual characteristics.

My results indicate that Medicaid increases the utilization of health care for preventive purposes in the same period in which a child is eligible for Medicaid coverage, but only for children with relatively low family income (between 100% and 185% of the poverty line, which I will call the *"low income group"* hereafter). Medicaid does not induce higher preventive health care utilization for children with relatively high family income (between 185% and 250% of the poverty line, from now onwards I will call the *"high income group"*).

The results also suggest that the short run effects of Medicaid on children's health are null or even negative. In the medium run −between 1 and 5 years after being eligible for Medicaid coverage− I find that Medicaid still affects children's health outcomes. Furthermore, I find that the effects of Medicaid on children's health are heterogeneous across children with different family income levels. While Medicaid is likely to have some positive effects on the low income children's health, it is also likely to have some negative effects on the high income children's health.

One possible explanation for the heterogeneous effects of Medicaid on the health of children with different levels of family income is the heterogeneous impact of the program on preventive health care utilization. My findings provide evidence that utilization could be a channel explaining the results because Medicaid only increases preventive health care utilization (measured as whether the child has visited a doctor at least once in the last 12 months for a routine health check-up) of children in the low income group and not of children in the high income group. An improvement in preventive health care utilization may be the reason of the positive effect of Medicaid on low income children's health in the medium run. This explanation, however, is not enough to explain why Medicaid has some negative effects on the high income children's health.

I argue that Medicaid has some negative effects on children in the high income group because it may induce families in this group to drop private health insurance. This switch may imply a reduction in the quality of health care services children can have access to. If the quality of the private health insurance is a normal good, then higher income families are the ones facing the following trade-off: taking the public insurance saves them money by quitting their child's private insurance at the cost of losing health care quality for their child if their private insurance allowed for better care quality. In an independent and simultaneous work, Koch (2010) also finds evidence that supports this hypothesis.

Some previous studies address the question of whether health insurance has a positive effect on children's health. Among those analyzing Medicaid, the results are mixed. For example, Currie and Gruber (1996) find evidence that the expansions in Medicaid eligibility thresholds between 1984 and 1992 increased the utilization of medical care and reduced child mortality. In contrast, Currie et al. (2008) find that expansions in Medicaid eligibility thresholds from 1986 to 2005 had no contemporaneous effects on the health of children between 9 and 17 years old, as reported by their parents. Their estimates, however, suggest that expansions that affected children of ages between 2 and 4 are associated with better health by the time they are 9-17 years old.

There is also an extensive literature studying the extent to which Medicaid expansions have led eligible families to switch from the private to the public health insurance (Cutler and Gruber, 1996; Lo Sasso and Buchmueller, 2004; Card and Shore-Sheppard, 2004; Ham and Shore-Sheppard, 2005; Gruber and Simon, 2007; Koch, 2010)). None of these papers, except Koch (2010), addresses the consequences of this "crowding-out" effect on children's health.

This paper contributes to the literature in several ways. First, I analyze both the contemporaneous and the lagged effects of Medicaid on different measures of health. The paper by Currie et al. (2008) is among the first to attempt estimating these lagged effects. However, in the cross sectional datasets they use, they must impute the family income and the state of residence of the child, since these variables are not observed during childhood. In contrast, I exploit the panel dimension of PSID data to match past eligibility with current health outcomes. Second, the identification strategy I propose allows for the estimation of Medicaid effects that vary across different levels of income. Results show the importance of this disaggregation when drawing any conclusions about the effects of the program. Finally, I propose an explanation for the existence of persistent negative effects of Medicaid on the high income group, suggesting that the "crowding-out" effect the public insurance generates may have a cost in terms of children's health, as long as there are quality differences between Medicaid and private insurances. I also bring an explanation for the persistent positive effects of Medicaid on the low income group through the utilization channel.

The remainder of the paper is organized as follows: Section 2 describes the Medicaid program; Section 3 presents the empirical strategy; Section 4 describes the data; Section 5 validates the regression discontinuity strategy; Section 6 presents and discusses the results; and Section 7 concludes.

2 Medicaid Program

The Medicaid program was introduced in the late 1960s as a health insurance component for state cash welfare programs targeting low-income single female head families. Medicaid is jointly financed by the federal government and the states. The federal government matches state spending on Medicaid.[2] The program is administered by the states and each state sets its own guidelines regarding eligibility and services, but subject to federal rules requiring minimum levels of coverage and services.

Medicaid eligibility for children was in its origins tied to the participation in the Aid for Families with Dependent Children (AFDC) program. Since the mid 1980s the linkage between AFDC coverage and eligibility for Medicaid has been gradually weakened, by eliminating the family structure requirements for young children and by allowing states to increase the income

[2] The federal share of Medicaid spending is determined by the Federal Medical Assistance Percentage (FMAP), which varies by state based on state per capita income relative to national average (Kaiser Commission on Medicaid and the Uninsured, 2010).

thresholds that determine eligibility (Currie and Gruber, 1996). The increase in the thresholds was first a state option, but later minimum levels of coverage were imposed by federal mandates. By April 1990, states were required to offer coverage to all children under 6 years old in families with income up to 133% of the poverty line and, starting in July 1991, they were required to provide coverage to all children under age 19, who were born after September 1983 and lived in households with incomes below 100% of the poverty line. As a result, by the mid-1990s, most children in the US living in households with incomes below 100% of the poverty line, and all young children living in households with incomes below 133% of the poverty line were eligible for Medicaid.

In practice, most states opted to raise the income thresholds beyond 133% of the poverty level and some did further increases using own state funds. States also set different threshold levels for different age groups. In 1997, the Medicaid program for children was augmented by the Children's Health Insurance Program (CHIP), which provided extra funds to expand eligibility for children beyond the existing limits of the Medicaid program. The CHIP program was implemented either by expanding the Medicaid program, or designing a new program, with features that mimic private health insurance (Gruber and Simon, 2007).

State Medicaid programs must cover mandatory services specified in federal law in order to receive federal matching funds. Medicaid covers a very comprehensive set of benefits and services for children under 21, defined by the pediatric Medicaid benefit also known as Early and Periodic Screening, Diagnostic, and Treatment (EPSDT) (Kaiser Commission on Medicaid and the Uninsured, 2010). The type of services that Medicaid must cover for children according to the federal rules include screening, preventive, and early detection services.[3] Health care must be made available to correct or ameliorate defects and physical and mental illnesses or conditions discovered by the screening services. Children also have access to physician and hospital services (impatient and outpatient). These services are provided with little or no copayment required (Gruber and Simon, 2007).[4] In terms of the package of services covered, Medicaid tends to be more generous than many private insurance plans.

Medicaid buys services primarily in the private health care sector. States pay health care providers on behalf of the Medicaid beneficiaries. States may purchase services on a fee-for-service basis or by paying premiums to managed care organizations (Kaiser Commission on Medicaid and the Uninsured, 2010). States also determine the rules to reimburse health care providers. In most cases, Medicaid's reimbursement is lower than the obtained from private insurance, which may induce some physicians to reject Medicaid patients or to lower the quality of the service provided.[5]

[3]Screening services include all the following services: comprehensive health and developmental history, immunizations, laboratory tests, lead toxicity screening, vision services, dental services, and hearing services.

[4]Copayments for some services were allowed to be higher for those above 150% of the poverty line since 2005. Cost-sharing for preventive care is prohibited for children. Premiums were prohibited for children until 2005 and remain prohibited for children under 150% of the poverty line. However, for those above 150% the poverty line, premiums and cost sharing cannot exceed 20% of the cost of the service. Additionally, total premiums and copayments cannot exceed 5% of family income for any family (Kaiser Commission on Medicaid and the Uninsured, 2010).

[5]For example, Decker (2007) finds that higher Medicaid fees increase the number of private physicians, especially in medical and surgical specialties, who see Medicaid patients. She also finds that higher fees also lead to visit times with physicians that are more comparable to visit times with private pay patients. Another paper by

3 Empirical Research Design

3.1 Contemporaneous Effects

The main objective is to estimate a simple model of the causal effect of Medicaid coverage on children's health care utilization and health outcomes

$$y_{it} = \alpha + \beta M_{it} + u_{it}, \tag{1}$$

where y_{it} is child i's outcome (utilization or health) in period t and M_{it} indicates whether the child had Medicaid coverage that same period. A simple OLS regression of equation (1) would yield a biased estimate. Medicaid coverage is an endogenous variable, because the access to this type of coverage is correlated with family income. Even after controlling for family income, selection problems may still be present because Medicaid enrollment is not mandatory. Among eligibles, the decision to take Medicaid may be correlated with other unobserved characteristics that are correlated with the outcomes.

In order to identify the effect of interest, I exploit the rule of assignment into Medicaid that allows me to implement a Regression Discontinuity (RD) design. The RD design is a quasi-experimental design with the defining characteristic that the probability of receiving the treatment changes discontinuously as a function of the variable that determines eligibility, called the assignment or forcing variable (Hahn et al., 2001).[6]

The intuition behind the RD is the following. Assuming that the eligibility threshold is exogenously given and families have imperfect control over their income, the eligibility status of a child with family income in the neighborhood of the threshold is randomly assigned, i.e., the rule generates a "local" randomized experiment. Making the additional assumption that in the absence of the treatment the outcome is a smooth function of income, the causal effect of Medicaid eligibility can be identified by comparing the average outcome of children just below the income threshold ("treatment group") with that of children just above it ("control group"). Any difference observed between these two groups can be attributed to the availability of treatment for treatment group members. Since enrollment in Medicaid is not mandatory −i.e., the coverage indicator, M_i, is not necessary equal to an indicator of eligibility status, Eli_i, which takes the value one if the child is eligible for Medicaid− comparing outcomes of eligible and non eligible individuals close to the threshold identifies the average effect of assignment into treatment or the *intention to treat effect* (ITT) at the threshold.[7]

The ITT effect can be significantly lower in absolute value than the effect the program has on those who are actually covered by Medicaid. Under the assumptions that the probability of having Medicaid coverage as a function of income is discontinuous at the threshold and that,

Cunningham and O'Malley (2009) finds that not only reimbursement fees matters, but also delays in reimbursement. They find evidence that Medicaid reimbursement time affects physicians' willingness to accept Medicaid patients.

[6]For a comprehensive discussion of the RD design and its application in economics see Imbens and Lemieux (2008), van der Klaauw (2002), and Lee and Lemieux (2010)

[7]For instance, studies such as Currie and Gruber (1996) and Currie et al. (2008), although using different identification strategies than in this paper, identify the intent to treat effects of Medicaid on children who where newly eligible to receive Medicaid benefits with the Medicaid expansion.

in the absence of the treatment, the association between the outcome variable and income is smooth, the parameter β can be estimated using the eligibility indicator Eli_i −which is randomly assigned in the neighborhood of the threshold− as an instrument for Medicaid coverage. This is called a "fuzzy" RD design (Hahn et al., 2001; Imbens and Lemieux, 2008).[8]

Ideally, to identify the causal effect it would be sufficient to compare outcomes of individuals above and below the threshold, in a very narrow interval around it. In practice, however, this is sometimes not possible because only few observations close to the threshold are available in the dataset. To overcome this problem, I implement a parametric RD specification as proposed by van der Klaauw (2002), that controls for a flexible function of the family income −assignment variable. I estimate β by 2SLS, where I instrument the treatment dummy, M_i, with the eligibility status, Eli_i. I follow a similar functional specification as in Carneiro and Ginja (2009).

The two equation system is given by

$$y_{it} = \alpha + \beta M_{it} + k_{2g}(z_{it}; \alpha_{2g}) + u_{it}, \tag{3}$$

$$M_{it} = \pi_0 + \pi_1 Eli_{it} + k_{1g}(z_{it}; \alpha_{1g}) + v_{it}, \tag{4}$$

where $Eli_{it} = \mathbf{1}\{\frac{z_{it}}{PL_t} \leq T_t\}$, is a dummy variable that takes the value one if the child is eligible for Medicaid, i.e., when family income (z_{it}), as a percentage of the poverty line (PL_t), is below the eligibility threshold (T_t); $k_{1g}(.)$ and $k_{1g}(.)$ are polynomials of order g of family income and u_{it} and v_{it} are unobserved error components. The periods for which I observe the outcomes are t=1997, 2002, 2007, as I explain in Section 4. Since the model is exactly identified, 2SLS estimates of β are numerically identical to the ratio of the reduced form coefficients θ/π_1, provided the same order of polynomial is used for $k_1(.)$ and $k_2(.)$ (Lee and Lemieux, 2010). The baseline estimates use fourth order polynomials.

The parametric specification in equation (3) allows to retain observations that are not necessarily close to threshold. The polynomial function of income controls for variation in the outcome and participation coming from income differences far from the threshold. Hence, β captures differences in the outcome variable for individuals just at the threshold. To check the robustness of the results, I run the regressions narrowing the width of the interval in the neighborhood of the threshold.

Hahn et al. (2001) were the first to suggest estimating the treatment effect in the fuzzy RD setting using two-stage least-squares (2SLS). Furthermore, they also point out that the estimate of β can be interpreted as a *Local Average Treatment Effect* (LATE) at the threshold under the same assumptions as in Imbens and Angrist (1994). Under these assumptions, the LATE

[8] As shown by Hahn et al. (2001), the treatment effect can also be recovered by dividing the "jump" in the relationship between the outcome and eligibility −the ITT at the threshold− by the fraction of individuals induced to take Medicaid at the threshold

$$\beta = \frac{\lim_{z \to z_0^-} E[y_i|z_i = z] - \lim_{z \to z_0^+} E[y_i|z_i = z]}{\lim_{z \to z_0^-} E[M_i|z_i = z] - \lim_{z \to z_0^+} E[M_i|z_i = z]}, \tag{2}$$

where z_i is the family income and z_0 is the eligibility threshold.

Hahn et al. (2001) were the first to show the connection between how the treatment effect is defined in the fuzzy RD design and the estimation of the treatment effect in an instrumental variables setting, when the instrument is a binary variable.

is defined as the average effect of treatment on the population of *"compliers"*, those eligible individuals at the threshold who receive the treatment if and only if they are assigned to it.

Given that Medicaid is a state administered program and that each state sets its own eligibility threshold, there are multiple thresholds at a given point in time. Assuming heterogeneous effects at different family income levels, hence, at different eligibility thresholds, the estimates obtained using a sample that pools all thresholds would estimate an average of the LATE across thresholds.

I also estimate the reduced-form equation that recovers the IIT effects

$$y_{it} = \alpha + \theta Eli_{it} + f_g(z_{it}; \gamma_g) + u_{it}, \tag{5}$$

where $f_g(z_{it}; \gamma_g)$ is a flexible function of income, a fourth-order polynomial in my baseline estimations. The parameter θ captures the ITT effect at the threshold, and given that there is not perfect compliance, this parameter is always a lower bound of β.

3.2 Lagged Cumulative Effects

In order to estimate the medium run causal effects of Medicaid on children's health I also take advantage of the "local" random assignment that the eligibility rule generates in a period $t-\tau$ to estimate the effects that Medicaid has, τ periods later, on period t outcomes using the following specification

$$y_{it} = \alpha + \theta_\tau Eli_{i,t-\tau} + f_g(z_{i,t-\tau}; \gamma_\tau) + u_{it}, \tag{6}$$

where $Eli_{i,t-\tau}$ is a dummy variable that takes the value one if the child was eligible for Medicaid in period $t-\tau$ and f_g is a polynomial of order g of income in period $t-\tau$, $z_{i,t-\tau}$.[9]

The parameter θ_τ does not isolate the direct effect of eligibility in period $t-\tau$ on period t outcomes, because of the possibility of multi-treatment. That is, between periods $t-\tau$ and t a child may have multiple opportunities to be eligible and enrolled in Medicaid. To the extent that period $t-\tau$ eligibility affects posterior participation in Medicaid, then the parameter θ_τ will also capture the indirect effect that subsequent participation may have on health outcomes of period t.

Given the possibility of multi-treatment, the marginal effect of making a child randomly eligible for Medicaid in a period $t-\tau$ on health outcomes in period t reflects a cumulative effect which is the sum of: 1) a direct effect on health outcomes τ years later, if it were possible to prohibit the child from being assigned to treatment in any other subsequent period; 2) an indirect effect on health outcomes through the effects on subsequent participation in the program. The total effect or *medium run ITT effect* of Medicaid eligibility on subsequent health, captured by θ_τ, is the effect of exogenously making a child eligible in a given period, without controlling for

[9] For the medium run analysis I restrict to estimating the ITT effects given that these effects provide a lower bound of the average treatment effects and the IV estimates tend to be more imprecise.

the family behavior in subsequent years. Following Cellini et al. (2010) the ITT parameter is[10]

$$\theta_{\tau}^{ITT} = \frac{\mathbf{d}y_{it}}{\mathbf{d}Eli_{i,t-\tau}} = \frac{\partial y_{it}}{\partial M_{i,t-\tau}} \times \frac{\partial M_{i,t-\tau}}{\partial Eli_{i,t-\tau}} + \sum_{h=1}^{\tau} \left(\frac{\partial y_{it}}{\partial M_{i,t-\tau+h}} \times \frac{\partial M_{i,t-\tau+h}}{\partial Eli_{i,t-\tau}} \right), \tag{7}$$

where $\frac{\partial y_{it}}{\partial M_{i,t-\tau}}$ is the direct effect of Medicaid in period $t-\tau$ under the assumption that the child would not have access to Medicaid in the subsequent years, and $\frac{\partial M_{i,t-\tau+h}}{\partial E_{i,t-\tau}}$ is the effect that eligibility in period $t-\tau$ has on subsequent Medicaid participation.

4 Data

The datasets used in the analysis are the Panel Study of Income Dynamics (PSID) and the Child Development Study (CDS) supplement. The CDS is a sample of children who were between 1 and 12 years old by 1997 and it contains information about children's health care utilization and health outcomes, obtained from the children's primary caregiver, as well as characteristics such as age and race of the child. Data for this cohort of children were collected in three waves: 1997, 2002, and 2007. Information on family income, Medicaid coverage, and family characteristics comes from the PSID dataset which can be matched with the CDS. I use the three CDS waves-matched with PSID data as repeated cross-sections, and I restrict the sample to children between 5 and 18 years old in any of the three waves. I keep only children for whom I can keep track of their eligibility and Medicaid status up to 5 years before the outcomes are observed.

I assign the Medicaid eligibility status of each child in the survey on a yearly basis. To impute eligibility I compare the annual family income as a percentage of the poverty line with the corresponding eligibility threshold, that is

$$Eli_{it} = \mathbf{1}\{\frac{\text{income}_{it}}{PL_t(fasize_{it})} \leq T_t(\text{state}_{it}, \text{age}_{it}\}, \tag{8}$$

where PL_t is the federal poverty line in period t (a function of the family size), and $T_t(.)$ is the state-age specific threshold in period t. I use the family income and the annualized official poverty threshold provided in the PSID data file for each family.[11] I get the information of state-age-year specific threshold from various reports of the National Governors' Association.

I use three types of outcome variables: one measure of preventive health care utilization; two objective measures of health; and two subjective measures of health. The measure of preventive health care utilization is a variable that indicates whether the child had visited a doctor at least once in the last 12 months for a routine health check-up. This measure is generally used to capture the utilization of medical resources for preventive purposes.[12]

[10]The main difference with Cellini et al. (2010) is that in their paper they have a "sharp" RD design, that is, being eligible is equivalent to receiving the treatment.

[11]See Grieger et al. (2009) for further details on the measures of family income and poverty thresholds in PSID. All income measures are expressed in 2000 US dollars.

[12]Other measures of health care utilization, such as the number of hospitalizations, may confound access and morbidity, as pointed out by Currie and Gruber (1996). An absence of a doctor visits for a regular check-up, however, better reflects an "access" problem.

As an objective measure of health I use the Body Mass Index (BMI).[13] A child's weight status is determined based on an age and sex-specific percentile for BMI. A child is classified as obese if her BMI is at or above the 95th percentile of the BMI distribution of children of the same age and sex. A child is overweight if her BMI is at or above the 85th percentile but below the 95th percentile.[14] Medicaid coverage may facilitate and increase the contact with physicians, which in turn increases the likelihood that children's weight status is monitored. Physicians recommendations about the quality of the diet and the adequate level of physical activity may be critical inputs to improve children's health status.

Additionally, I use two subjective health measures, both reported by the child's caregiver: an indicator of whether the child has an excellent health status and a dummy variable for whether the child missed more than five days of school due to illness during the last 12 months. The first measure reflects the caregiver's perception about the child's overall health status. I interpret any deviation from excellent health as reflecting some health problem. The second measure links child's health status and school attendance, capturing a key aspect of how health may affect her human capital formation. If Medicaid allows to prevent illnesses it might also help to avoid missing school days.

One drawback of measuring the effects of Medicaid on subjective health measures is that these effects may be difficult to interpret. Currie and Gruber (1996) argue that these measures may capture two possible effects. If the public insurance coverage leads individuals to increase the contacts with the medical system, then there could be a "true" effect on child health, resulting in better child's health reports. The increased contacts with physicians, however, may also affect parents' perception about the health of the child. Parents may learn about health conditions the child already had but they were not aware of because they did not contact physicians so frequently before having the public insurance coverage. Also, if targeted children are switching from a private insurance to the public, parent's reports may be sensitive to perceived changes in the quality of health care they have access to with the public insurance instead of the private one.

Columns (1) and (2) of Table 1 present descriptive statistics of children's and family main characteristics, for the full sample. I refer as "full" sample as the sample pooling all eligibility thresholds. Here, I consider all children whose annual family income is within a distance of $\pm$ 50 thousand dollars in period t, for t=1997, 2002, and 2007, although for the empirical analysis I restrict to narrower intervals in the neighborhood of the threshold. Columns (3) to (8) present the same descriptives but for three subsamples, defined by the level of Medicaid "generosity" in each state, where the generosity is determined according to the level of the income threshold that determines eligibility. The first subsample consists of children living in states where the generosity of Medicaid coverage is relatively low (the eligibility thresholds are lower than 185% the poverty line); the second subsample includes children living in states with a middle level of generosity (the eligibility thresholds are set between 185% and 250% the poverty line); and

[13]Although it is not completely "objective", since during the interview, the primary caregiver reports the weight of the child, and the interviewer measures his or her height.

[14]The CDS dataset provides indicators of the child's obesity and overweight status according to this definition, based on the Centers for Disease Control and Prevention (CDC) growth charts. Each of the CDC BMI-for-age gender specific charts contains a series of curved lines indicating specific percentiles. See the CDC Growth Charts for children at: http:\www.cdc.gov\growthcharts.

finally, the third subsample consist of children living in states with relatively high levels of generosity (the eligibility thresholds are above 250% the poverty line).

From columns (1) and (2) it is clear that Medicaid eligibles are more disadvantaged than non-eligibles in several dimensions. They have lower family income −by definition of eligible−, they are more likely to be minorities, to live in a female-headed family, and to live with a less educated head of household. They are worse off in terms of health outcomes. However, they are more likely to have visited a doctor for a check up in the last 12 months. A similar pattern emerges if I split the sample according to the different levels of Medicaid's coverage generosity. In the three groups, eligible children are always more disadvantaged in terms of socioeconomic characteristics, they tend to have worse health outcomes, and to use more preventative health are services. The only exception is the states with higher levels of generosity, where utilization is higher for non-eligibles.

Only 53% of eligible children are actually enrolled in Medicaid, although enrollment is heterogeneous depending on family income level.[15] The incentives to enroll in Medicaid decrease with income, as it can be observed by comparing eligible children in states with higher levels of Medicaid generosity. The take-up rate is 61% in states with modest Medicaid coverage generosity, where eligibles' average family income is 12.1 thousand dollars per year. This proportion falls to 53% in states with middle level generosity and where eligibles' average income is 21.4 thousand dollars, and it is even lower (35%) in states with the most generous coverage, where eligibles' average income is 33,6 thousand dollars. The incentives to enroll in Medicaid may decline with income because, as income rises, the family's financial constraint is less binding, which allows them to acquire an alternative source of coverage in private markets.

[15] Note that among non-eligibles there are individuals with Medicaid coverage. This happens because there may be timing problems in the reports of individuals family income −from which I infer eligibility status− and Medicaid coverage. Also, income fluctuations during the year can make an individual eligible for Medicaid at some point of the year but according to the annual income I they re clasified as non-eligible. Approximately 10% of the non-eligibles in the full sample report having Medicaid, although this percentage rise up to a 20% for the subgroup of individuals just above the eligibility threshold, as will be showed in Section 5.

Table 1: Descriptive Statistics

	Full sample		Thresholds<185% PL		Thresholds [185,250]% PL		Thresholds>250% PL	
	Eligible (1)	Non-Eligible (2)	Eligible (3)	Non-Eligible (4)	Eligible (5)	Non-Eligible (6)	Eligible (7)	Non-Eligible (8)
Outcome Measures								
Visited a doctor at least once in last 12 months	0.75	0.68	0.70	0.62	0.76	0.71	0.74	0.77
Obese (over 95 percentile BMI dist.)	0.24	0.19	0.23	0.19	0.24	0.20	0.28	0.17
Overweight (85-95 percentile BMI dist.)	0.15	0.15	0.15	0.15	0.15	0.16	0.14	0.18
Obese + Overweight	0.39	0.35	0.38	0.34	0.39	0.36	0.42	0.35
More than 5 days of school missed	0.12	0.12	0.18	0.12	0.10	0.13	0.12	0.11
Health Excellent	0.39	0.54	0.27	0.47	0.42	0.58	0.47	0.62
Insurance Coverage								
Medicaid Coverage	0.53	0.09	0.61	0.12	0.53	0.07	0.35	0.05
Private Insurance	0.32	0.81	0.21	0.77	0.32	0.82	0.58	0.90
Family and Child Characteristics								
Family Income (2000 dollars)	20,438	53,349	12,101	44,120	21,352	58,124	33,641	78,361
	(12,481)	(18,574)	(7,325)	(15,010)	(11,076)	(16,290)	(15,220)	(20,573)
Income Cutoff (eligibility threshold)	35,768	30,898	22,465	21,315	37,260	35,681	56,647	58,971
	(13,557)	(12,832)	(6,755)	(5,993)	(9,339)	(8,384)	(13,475)	(14,546)
Income threshold as % of poverty line	200.65	174.30	123.68	120.02	205.67	201.91	341.23	331.43
	(66.33)	(61.07)	(24.56)	(24.62)	(18.05)	(16.45)	(46.78)	(41.84)
Metropolitan Area	0.67	0.60	0.76	0.60	0.64	0.57	0.61	0.80
Rural Area	0.15	0.16	0.11	0.14	0.15	0.20	0.20	0.09
	0.35	0.37	0.31	0.35	0.36	0.40	0.40	0.28
Family Size	4.20	4.09	4.30	4.09	4.24	4.08	3.75	4.08
	(1.35)	(1.13)	(1.34)	(1.07)	(1.38)	(1.19)	(1.16)	(1.16)
Education (yrs.) of the Head of the Household	11.84	12.92	11.91	12.66	11.69	12.97	12.55	14.10
	(2.29)	(2.04)	(1.84)	(2.01)	(2.47)	(2.00)	(1.98)	(2.13)
Female Head	0.61	0.23	0.68	0.27	0.60	0.21	0.46	0.16
Child Age	11.20	10.97	9.41	9.57	11.82	12.09	11.74	12.30
	(3.17)	(3.28)	(3.18)	(2.96)	(2.93)	(3.05)	(2.95)	(3.26)
Male	0.52	0.50	0.49	0.50	0.52	0.51	0.58	0.52
Black	0.61	0.35	0.73	0.45	0.61	0.29	0.37	0.30
Hispanic	0.06	0.03	0.00	0.01	0.09	0.04	0.00	0.00
Birth Weight (kg)	3.20	3.36	3.17	3.34	3.19	3.37	3.30	3.38
	(0.69)	(0.64)	(0.72)	(0.65)	(0.67)	(0.64)	(0.71)	(0.60)
Mother Age at Child's Birth	25.16	27.02	25.22	26.61	25.06	27.21	25.63	27.94
	(6.01)	(5.37)	(5.97)	(5.14)	(6.12)	(5.55)	(5.49)	(5.29)
N	1,127	1,668	284	700	713	835	130	105

Notes: Observations are restricted to children in the CDS whose family income is at a distance of ± 50 thousand dollars from the threshold in years 1997, 2002 or 2007. Columns (1) and (2) present descriptive statistics for the full sample. Columns (3) and (4) correspond to the subsample of children living in states where the generosity of Medicaid coverage is relatively low −the eligibility thresholds are lower than 185% of the poverty line; Columns (5) and (6) correspond to the subsample of children living in states with a middle level of generosity −the eligibility thresholds are set between 185% and 250% of the poverty line; Columns (7) and (8) correspond to the subsample of children living in states with relatively high levels of generosity −the eligibility thresholds are above 250% of the poverty line.

5 Validity of RD Design: Robustness Analysis

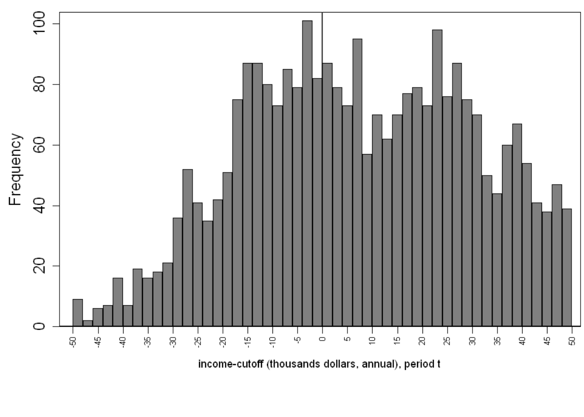
A first step in the analysis involves testing the identification assumptions of the RD to check its internal validity. The empirical strategy is based on the assumption that eligibility to receive Medicaid coverage is as good as randomly assigned in the neighborhood of the income thresholds. This assumption requires families to be unable to manipulate their incomes perfectly well so that they cannot control if their children qualify for Medicaid. Additionally, for the validity of the design, the probability of participating in Medicaid as a function of family income should show a discontinuity at the threshold. Finally, an implication of the "local" randomization is that individuals at either side of the threshold should be similar both in observed and unobserved characteristics.

To check the validity of the RD design I perform the following robustness analysis, as proposed by Imbens and Lemieux (2008) and Lee and Lemieux (2010). First, I inspect the histogram of the family income −the assignment variable− to check whether families have imprecise control over it. A spike to the left of the threshold may indicate that families are manipulating their income to fall below the eligibility threshold. Second, I estimate the participation equation to check whether Medicaid eligibility rule induces a discontinuity at the threshold. Finally, I examine whether baseline covariates (variables that should not be affected by the program as well as individual characteristics not taken into account to determine eligibility) are balanced on either side of the threshold.

5.1 Manipulation of the assignment variable

Figure 1 presents an histogram with the distribution of family income, pooling all observations for the years 1997, 2002, and 2007. This graph shows the number of observations within bins of 2 thousand dollars width. Given that there are multiple thresholds, income is normalized by subtracting the corresponding eligibility threshold. A negative value indicates that income is below the threshold and the child is eligible for Medicaid. An accumulation of observations below the normalized threshold (equal to zero) may be an indication of income manipulation. At first sight families do not seem to be manipulating their income in order to be below the threshold. There are spikes both to the right and to the left of the threshold.

Figure 1: Family Income Distribution. Full sample. Years 1997, 2002 and 2007.



McCrary (2008) proposes a simple two-step procedure for testing whether there is a discontinuity in the density of the assignment variable. Implementing the McCrary test on this sample, I reject the hypothesis of a discontinuity of the density function of income at the threshold. Results of this test are reported in Appendix A.[16]

5.2 Discontinuity in the probability of participating in Medicaid

As discussed in Section 3, despite of imperfect compliance, the fuzzy RD analysis can identify a LATE at the threshold as long as the eligibility rule generates a jump in the participation rate at the threshold.

Figure 2 plots the proportion of children who are enrolled in Medicaid over family income for the years 1997, 2002, and 2007. Each dot is the proportion of children with Medicaid coverage within a family income bin of 2 thousand dollars width. The solid lines are predictions from local linear regressions with bandwidth of 5 thousand dollars estimated with the raw data. We can observe that at the threshold −normalized to 0− the probability of participation has a discontinuity of about 15 percentage points.

Figure 2: Medicaid Participation. Full sample. Years 1997, 2002 and 2007.

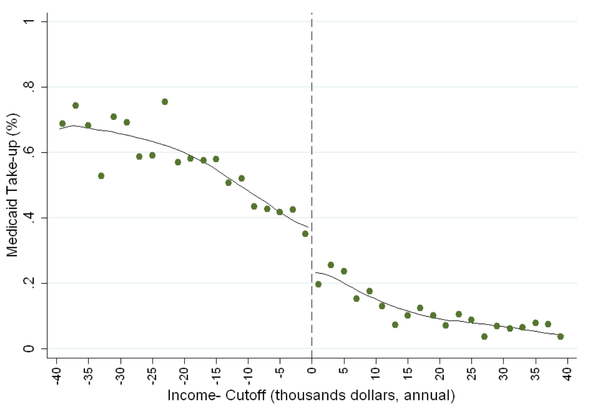


Table 2 reports the results of the parametric estimation of the participation equation specified as

$$M_{it} = \pi_0 + \pi_1 Eli_{it} + k_{1g}(z_{it}; \alpha_{1g}) + u_{it}, \qquad t = 1997, 2002, 2007, \tag{9}$$

where Eli_{it} is the eligibility indicator in period t, M_{it} is Medicaid enrollment status in the same period, and $k_{1g}(.)$ is a polynomial of order g of family income, z_i.

Panel A of Table 2 reports the estimated jump in the probability of participation at the threshold −pooling all thresholds. As discussed in Section 3.1, because of the small sample sizes I would get by restricting the analysis only to observations in a tight neighborhood of the threshold, each column of this table shows the estimates of the same model but considering windows of different widths around the threshold. The results indicate that making a child

[16] To check the robustness of the results, I perform the same exercise on a sample that considers all the years for which I can track the family income in PSID for the children in my sample (1991-2007). Using this extended sample I also reject the null hypothesis of a discontinuity of the density distribution at the threshold. Results available upon request.

with family income equal to the threshold eligible for Medicaid increases the probability of enrollment by between 14 and 20 percentage points, depending on the width of the interval around the threshold.

Panel B of Table 2 reports the estimates for the discontinuity at the threshold, but allowing for the heterogeneous jumps depending on the threshold level. The model estimated is the following

$$M_{it} = \gamma_0 + \sum_{j=1}^{2} \gamma_j T_{j,it} + \sum_{j=0}^{2} \pi_j Eli_{j,it} + k_{0g}(z_{it}; \alpha_{0g}) + \sum_{j=1}^{2}, k_{jg}(z_{it}; \alpha_{jg}) \times T_{j,it} + u_{it}, \qquad (10)$$

where $T_{j,it}$ is an indicator that takes the value one if child i lives in period t in a state where the eligibility threshold is $T_j\%$ of the poverty line, and $Eli_{j,it} = Eli_{it} \times T_{j,it}$, is an indicator that takes the value one if the child is eligible for Medicaid and lives in a state where the eligibility threshold is $T_j\%$. I consider three categories of T: thresholds lower than 185% of the poverty line (baseline category, T_0), thresholds between 185% and 250% of the poverty line (T_1), and thresholds higher than 250% of the poverty line (T_2).

Table 2: Participation Equation. "Jump" at the threshold. Years 1997, 2002, 2007.

	Bandwidth (thousands dollars)			
	±30	±20	±15	±2
A. Full sample				
Eli$_t$	0.143***	0.158***	0.163***	0.201***
	(0.036)	(0.040)	(0.042)	(0.076)
B. Model Interacted				
Eli$_t$ $\times 1\{T < 185\}$	0.263***	0.276***	0.258***	0.360**
	(0.063)	(0.071)	(0.073)	(0.149)
Eli$_t$ $\times 1\{185 \leq T \leq 250\}$	0.159***	0.222***	0.198***	0.224**
	(0.052)	(0.061)	(0.064)	(0.107)
Eli$_t$ $\times 1\{T > 250\}$	-0.035	-0.019	-0.022	-
	(0.060)	(0.063)	(0.066)	
N	2163	1555	1185	156

Notes: Robust standard errors (in parenthesis) are clustered at the family level. All regressions are linear probability models and all include a polynomial of order 4 of the log income, age, and family size, and year and state dummies. In each column the sample is restricted to observations with family income levels that fall within the indicated bandwidth. **Panel A:** Estimates in each column come from a separate linear probability model $M_{it} = \pi_0 + \pi_1 Eli_{it} + k_{1g}(z_{it}; \alpha_{1g}) + u_{it}$. **Panel B:** Estimates in each column come from a separate linear probability model $M_{it} = \gamma_0 + \sum_{j=1}^{2} \gamma_j T_{j,it} + \sum_{j=0}^{2} \pi_j Eli_{j,it} + k_{0g}(z_{it}; \alpha_{0g}) + \sum_{j=1}^{2} k_{jg}(z_{it}; \alpha_{jg}) \times T_{j,it} + u_{it}$.

The parameters π_j capture the jump in participation at the threshold in each of the three groups of states. The results show that the discontinuity in Medicaid participation is larger in states with the lower eligibility thresholds. In states where $T < 185\%$ the jump is between 26 and 36 percentage points, while in states where $185\% \leq T \leq 250\%$ it is between 16 and 22 percentage points. In states where the thresholds are above 250% the poverty line I do not find evidence that making a child eligible for Medicaid increases the chance that she receives Medicaid coverage. Higher-income families, targeted by Medicaid in more generous states, may not find beneficial to enroll their children in Medicaid because they may have better options available. This result is consistent with the quality of private health insurance being a normal good.[17]

Additionally, I take advantage of the panel structure of my dataset to perform a placebo test to check whether Medicaid participation in a period $t-\tau$ as a function of income in period t changes discontinuously at period t thresholds. If eligibility in period t is truly exogenous in the neighborhood of the threshold, then the only variable that should change discontinuously as a function of income in period t is Medicaid coverage in period t. Although there could be some correlation between income in period t and Medicaid participation in a period $t-\tau$ (because income is serially correlated), I should not observe any discontinuity in period $t-\tau$ participation at the period t threshold (i.e., eligibility in the neighborhood of the threshold in period t is exogenous and does not depend on previous Medicaid participation.) Since Medicaid participation across periods can be highly correlated, finding a discontinuity in participation in period t but not in $t-\tau$ would be a strong piece of evidence supporting the validity of the RD design (Lee and Lemieux, 2010).

In Section B.1 of Appendix B, I present graphs showing the relation between Medicaid coverage in period $t-\tau$ and family income in period t. The graphs show that participation in $t-\tau$ is negatively correlated with income in t and that it is a smooth function of family income in period t . Medicaid participation in periods $t-2$ and $t-3$ does not change discontinuously at the threshold although for period $t-1$, there is a small jump at the threshold. This is likely to happen because, on the one hand, most states guarantees a minimum of six months to one year of permanence in Medicaid with independence of their family income and, on the other hand, because family income may not change substantially from one year to the other.

5.3 Balance of individual characteristics on either side of the thresholds

The third robustness analysis consists on checking whether children characteristics are "locally" balanced, which is an implication of the "local" randomization generated by the eligibility rule. To check for this, I run regressions of the form

$$y_{it} = \gamma_0 + \gamma_1 Eli_{it} + f_g(z_{it};\gamma) + u_{it}, \qquad t = 1997, 2002, 2007, \tag{11}$$

where y_{it} are child and family characteristics not taken into account at the moment of determining Medicaid eligibility. I also consider pre-treatment variables which should not be affected by eligibility status, such as child's birth weight or mother's age at child's birth. If any of the observable characteristics changes discontinuously at the threshold, it is an indication that the eligibility rule does not generates a "local" randomization.

Table 3 presents the results and there are no signs of systematic discontinuous changes of characteristics at the threshold.

[17] These results remain the same when considering the sample that includes all years for which I can keep track family income in PSID (period 1991 and 2007). See Appendix B.

Table 3: Balance of covariates on either side of the threshold. Full sample. Years 1997, 2002, 2007.

	Bandwidth (thousands dollars)			
Dep. Var.	±30	±20	±15	±2
Male	0.083**	0.068	0.069	0.071
	(0.038)	(0.042)	(0.045)	(0.083)
Black	0.016	0.025	0.018	0.118
	(0.033)	(0.036)	(0.038)	(0.087)
Metropolitan Area	0.044	0.066	0.061	0.107
	(0.040)	(0.042)	(0.044)	(0.086)
Rural Area	-0.034	-0.047	-0.037	-0.062
	(0.035)	(0.038)	(0.039)	(0.060)
Child Birth Weight	-0.019	-0.039	-0.047	-0.091
	(0.059)	(0.062)	(0.068)	(0.132)
Head Education (yrs)	0.047	0.095	0.057	0.682*
	(0.181)	(0.189)	(0.205)	(0.368)
Mother age at child birth	0.481	0.436	0.523	0.218
	(0.493)	(0.527)	(0.555)	(1.061)
N	2163	1555	1185	176

Notes: Robust standard errors (in parentheses) are clustered at the family level. Each entry comes from a separate linear regression, $y_{it} = \gamma_0 + \gamma_1 Eli_{it} + f_g(z_{it}; \gamma_2) + u_{it}$, where the dependent variable is replaced by children and family characteristics, and pre-treatment covariates. The reported coefficient is $\hat{\gamma_1}$ of equation (11). Each regression includes 4th order polynomial of log of income, age, and family size as well as year and state dummies.

6 Results

6.1 Contemporaneous Effects

Preventive health care utilization. Table 4 presents the results of the contemporaneous effects of Medicaid −equations (3) and (5)− on utilization of preventive medical care for the full sample (pooling all eligibility thresholds), therefore, the effects reported in this table are average effects across thresholds. The intention to treat estimates show that making a child, with family income equal to the threshold, eligible for Medicaid slightly increases health care utilization by 5 percentage points relative to a similar child but non-eligible for Medicaid. The IV estimates, however, indicate that the average effect for the subpopulation of compliers −those who, made eligible for Medicaid, would enroll into the program− is between 30 and 35 percentage points. These estimates are, however, not statistically significant.

Table 5 reports the effects of Medicaid on utilization, but allowing for heterogeneous effects depending on the level of the eligibility threshold: states with thresholds under 185% of poverty line (the *"low income" group*), and states with thresholds between 185% and 250% (the *"high income" group*).[18] Panel A indicates that Medicaid eligibility induces around 14-17 percentage points increase in utilization for children between 5 and 18 years old in the low income group, with an average effect of about 46 to 52 percentage point for the compliers. For children in the high income group, Medicaid eligibility does not have a statistically significant impact on

[18]The effects in states with thresholds above 250% are not estimated because, as I showed in Section 5.2, Medicaid eligibility does not predict a jump in participation for these thresholds.

utilization and the coefficients are close to zero.[19]

Table 4: Contemporaneous effects of Medicaid on utilization. Children between 5 and 18 years old. Years 1997, 2002, and 2007. *Dep. Var.: The child has visited a doctor for a routine health check-up in the last 12 months.*

	Bandwidth (thousands dollars)		
Full sample	±30	±20	±15
Intention to treat			
Eli_t	0.048	0.053	0.047
	(0.037)	(0.041)	(0.044)
Outcome equation- IV-RD			
M_t	0.338	0.351	0.292
	(0.266)	(0.262)	(0.277)
N	2163	1555	1185

Notes: Robust standard errors (in parenthesis) are clustered at the family level. All regressions are linear probability models and all include a polynomial of order 4 of log income, age, and family size, year and state dummies. In each column the sample is restricted to observations with family income levels that falls within the bandwidth indicated. The intention to treat estimates in each column come from the following model: $y_{it} = \alpha + \theta Eli_{it} + f_g(z_{it}; \gamma_g) + u_{it}$. The IV-RD estimates in each column come from the following model: $y_{it} = \alpha + \beta M_{it} + k_{2g}(z_{it}; \alpha_{1g}) + u_{it}$, where eligibility instruments for Medicaid coverage.

Table 5: Contemporaneous effects of Medicaid on utilization. Children between 5 and 18 years old. Years 1997, 2002, and 2007. *Dep. Var.: The child has visited a doctor for a routine health check-up in the last 12 months.*

	A All Ages (5-18) Bandwidth (thousands dollars)			B Age group 5-11 Bandwidth (thousands dollars)			C Age group 12-18 Bandwidth (thousands dollars)		
Model Interacted	±30	±20	±15	±30	±20	±15	±30	±20	±15
Intention to treat									
$\text{Eli}_t \times 1\{T < 185\}$	0.138**	0.157**	0.169**	0.160**	0.202**	0.170*	0.087	0.119	0.148
	(0.065)	(0.072)	(0.076)	(0.077)	(0.086)	(0.092)	(0.126)	(0.127)	(0.141)
$\text{Eli}_t \times 1\{185 \leq T \leq 250\}$	-0.003	-0.005	-0.022	-0.075	-0.074	-0.142	-0.002	-0.047	0.029
	(0.050)	(0.060)	(0.065)	(0.095)	(0.105)	(0.109)	(0.062)	(0.080)	(0.089)
Outcome equation. IV-RD									
$\text{M}_t \times 1\{T < 185\}$	0.455*	0.524**	0.517*	0.424	0.668**	0.554	0.341	0.416	0.425
	(0.233)	(0.225)	(0.300)	(0.273)	(0.294)	(0.348)	(0.437)	(0.493)	(0.610)
$\text{M}_t \times 1\{185 \leq T \leq 250\}$	0.080	0.082	-0.071	-0.466	0.117	-0.432	0.181	-0.705	-0.086
	(0.263)	(0.237)	(0.291)	(0.714)	(0.355)	(0.429)	(0.390)	(0.663)	(0.443)
N	1992	1441	1102	1089	784	599	801	581	442

Notes: Robust standard errors (in parenthesis) are clustered at the family level. All regressions are linear probability models and all include a polynomial of order 4 of log income, age, and family size, year and state dummies. In each column the sample is restricted to observations with family income levels that falls within the bandwidth indicated. The intention to treat estimates in each column come from the following model: $y_{it} = \alpha_0 + \theta_0 Eli_{0it} + \theta_1 Eli_{1it} + f_{0g}(z_{it}; \gamma_{0g}) + \alpha_1 T_{1it} + f_{1g}(z_{it}; \gamma_{1g}) \times T_{1it} + u_{it}$. The IV-RD estimates in each column come from the following model: $y_{it} = \alpha_0 + \beta_0 M_{0it} + \beta_1 M_{1it} + \alpha_1 T_{1it} + k_{0g}(z_{it}; \gamma_{0g}) + k_{1g}(z_{it}; \gamma_{1g}) \times T_{1it} + u_{it}$, where eligibility instruments for Medicaid coverage.

Finally, panel B and C of Table 5 show the estimated effects of Medicaid on preventive health

[19] In Appendix C I estimate equations (3) and (5) considering different orders of polynomials. The sensitivity analysis shows that the estimates are robust to the model specification. See Table 12

care utilization for children of different age groups (from 5 to 11 and from 12 to 18). Larger and significant effects are observed for low income children of ages between 5 and 11 years old. Medicaid also positively affects the utilization of preventive medical care of children between 12 and 18 years old in the low income group, although the magnitude is lower than that of the younger group and not statistically significant.

Table 6: Contemporaneous effects of Medicaid on children's health outcomes. Children between 5 and 18 years old. Years 1997, 2002, and 2007.

	Bandwidth (thousands dollars)					
	±30	±20	±15	±30	±20	±15
	A. Excellent Health			**B. Obese**		
Intention to treat						
Eli$_t$ $\times 1\{T < 185\}$	-0.090	-0.085	-0.067	-0.055	-0.051	-0.016
	(0.066)	(0.075)	(0.082)	(0.060)	(0.068)	(0.069)
Eli$_t$ $\times 1\{185 \leq T \leq 250\}$	-0.093	-0.157**	-0.160**	-0.012	0.001	0.033
	(0.060)	(0.069)	(0.074)	(0.049)	(0.057)	(0.061)
Outcome equation						
IV-RD						
M$_t$ $\times 1\{T < 185\}$	-0.777	-0.574	-0.809	-0.391	-0.399	-0.141
	(0.485)	(0.463)	(0.665)	(0.374)	(0.389)	(0.433)
M$_t$ $\times 1\{185 \leq T \leq 250\}$	-0.877*	-0.816*	-0.992	-0.214	-0.184	0.195
	(0.529)	(0.489)	(0.784)	(0.371)	(0.332)	(0.469)
	C. Overweight			**D. School days missed**		
Intention to treat						
Eli$_t$ $\times 1\{T < 185\}$	0.055	0.035	0.001	-0.003	0.031	0.029
	(0.050)	(0.055)	(0.058)	(0.048)	(0.056)	(0.062)
Eli$_t$ $\times 1\{185 \leq T \leq 250\}$	0.018	0.035	0.013	-0.067*	-0.070	-0.060
	(0.040)	(0.048)	(0.053)	(0.038)	(0.044)	(0.046)
Outcome equation						
IV-RD						
M$_t$ $\times 1\{T < 185\}$	0.403	0.209	0.305	-0.173	0.162	0.202
	(0.301)	(0.289)	(0.405)	(0.290)	(0.311)	(0.407)
M$_t$ $\times 1\{185 \leq T \leq 250\}$	0.253	0.188	0.337	-0.384	-0.189	-0.228
	(0.302)	(0.263)	(0.461)	(0.312)	(0.280)	(0.394)
N	1993	1431	1101	1993	1431	1101

Notes: Robust standard errors (in parenthesis) are clustered at the family level. All regressions are linear probability models and include a polynomial of order 4 of log income, age, and family size, year and state dummies. In each column the sample is restricted to observations with family income levels that falls within the bandwidth indicated. The intention to treat estimates in each column come from the following model:
$y_{it} = \alpha_0 + \theta_0 Eli_{0it} + \theta_1 Eli_{1it} + f_{0g}(z_{it}; \gamma_{0g}) + \alpha_1 T_{1it} + f_{1g}(z_{it}; \gamma_{1g}) \times T_{1it} + u_{it}$. The IV-RD estimates in each column come from the following model: $y_{it} = \alpha_0 + \beta_0 M_{0it} + \beta_1 M_{1it} + \alpha_1 T_{1it} + k_{0g}(z_{it}; \gamma_{0g}) + k_{1g}(z_{it}; \gamma_{1g}) \times T_{1it} + u_{it}$, where eligibility instruments for Medicaid coverage.

Health outcomes. Table 6 presents the estimated contemporaneous effects of Medicaid on four measures of children's health: overweight, obesity, an indicator of excellent health, and an indicator of missing more than 5 school days due to illness. According to these results, Medicaid does not seem to have a positive effect on health in the short run for children between 5 and 18 years old. Moreover, Panel A shows that Medicaid has a negative impact on the probability of being in excellent health for children in the high income group. Since it is a subjective measure reported by children's caregivers it can be argued that this effect is just a "perception" effect and it does not reflect a real change in children's health. Medicaid just induces more contacts with physicians and parents become more aware of certain health problems their children already had. According to previous results on utilization (Table 5), this explanation may be only plausible to explain the negative effects of Medicaid on the probability of being in excellent health for children in the low income group (first row of Table 6), because Medicaid increases its preventive health care utilization. However, there is no evidence that Medicaid increases preventive health care utilization of children in the high income group. Hence, the perception effect does not explain why Medicaid has a negative and statistically significant effect on the probability of being in excellent health of this group, and, moreover, neither why this effect is even larger in magnitude relative to that on the same health measure of the low income group.

Other reason why parents of children covered by Medicaid in the high income group are more likely to report that their children are in worse health is because Medicaid may induce them to drop a private health insurance. If they perceive that Medicaid quality is lower than their previous private option, they may translate this perception to a worse evaluation of their children's health. I discuss this hypothesis in Section 6.3.

6.2 Lagged Effects on Health

The finding that Medicaid eligibility and coverage have no statistically significant contemporaneous effect on health of some subgroups may merely reflect that health is a stock and that the potential positive effects of the program are only visible after some time. Also, the finding that Medicaid negatively effects some subgroups may reflect possible "perception" effects on parents that are not really related to changes in children's health.

Now I turn to the analysis of the cumulative effects of Medicaid in the medium run. If the effects of Medicaid in the short run only capture "perception" effects then they should vanish with time, and they should not appear in regressions of health outcomes on past eligibility. On the contrary, if negative or positive effects persist after several periods, it is more likely that these effects are real effects on children's health.

Tables 7 and 8 report the cumulative IIT estimates, which capture the effect of making a child randomly eligible for Medicaid in a given period on the probability of being in excellent health and obesity after τ periods −equation (6).[20,21] These ITT estimates identify the effects of eligibility in one moment of time on future outcomes, without controlling for behavioral changes between the period of eligibility and the period in which outcomes are measured. Thus, IIT estimates reflect accumulated effects as shown in equation (7).[22]

Excellent Health. Table 7 reports the lagged ITT cumulative effects of Medicaid eligibility on the probability of having excellent health after τ periods, for the low −columns (1) to (3)− and the high income groups −columns (4) to (6). Column (5) shows that Medicaid has persistent, negative, and statistically significant effects after two periods on the health of children between 5 and 11 years old in the high income group. This result indicates that an eligible child of the

[20] I do not find significant lagged effects on the probability of being overweight and the probability of missing school days due to illness, nor in the probability of visiting a doctor for preventive purposes.

[21] I also estimate equation (6) considering different orders of polynomials and different intervals in the neighborhood of the threshold. Results available upon request.

[22] I do not report the IV estimates because they tend to be more imprecise. However, ITT effects are lower bounds for the average treatment effects.

high income group in a period t is, two years later, 18 percentage points less likely to be in excellent health than a similar non-eligible child.

On the other hand, column (3) shows that after three periods onwards Medicaid eligibility has a positive and statistically significant effect on the health of children between 12 and 18 years old in the low income group. That is, an eligible child of the low income group in a period t is, three years later, 19 percentage points more likely to be in excellent health than a similar non-eligible child. He is still 16 percentage points more likely to be in excellent health after 5 years.

Table 7: Lagged cumulative effects of Medicaid Eligibility on children's health. *Dep. Var: The child has Excellent Health.*

	Thresholds <185% ($\text{Eli}_t \times 1\{T < 185\}$)			Thresholds [185-250]% ($\text{Eli}_t \times 1\{185 \leq T \leq 250\}$)		
Time Elapsed	5-18 years old (1)	5-11 years old (2)	12-18 years old (3)	5-18 years old (4)	5-11 years old (5)	12-18 years old (6)
1 year (θ_1)	-0.063	-0.038	-0.092	-0.063	-0.045	-0.111
	(0.066)	(0.083)	(0.115)	(0.061)	(0.089)	(0.085)
2 years (θ_2)	-0.046	-0.061	-0.042	-0.085	-0.180*	0.001
	(0.061)	(0.079)	(0.110)	(0.065)	(0.094)	(0.087)
3 years (θ_3)	0.005	-0.100	0.193**	-0.005	-0.063	0.026
	(0.059)	(0.074)	(0.095)	(0.072)	(0.097)	(0.094)
4 years (θ_4)	0.071	0.029	0.149	0.001	0.029	0.010
	(0.063)	(0.079)	(0.092)	(0.072)	(0.101)	(0.096)
5 years (θ_5)	-0.020	-0.078	0.158*	-0.033	-0.070	0.005
	(0.054)	(0.069)	(0.087)	(0.084)	(0.111)	(0.112)

Notes: Robust standard errors (in parenthesis) are clustered at the family level. All regressions are linear probability models and all include a polynomial of order 4 of log income in period $t - \tau$, age, and family size, year and state dummies. The sample is restricted to observations with family income levels that falls within the bandwidth of $\pm$ 30 thousand dollars from the threshold in period $t - \tau$. The intention to treat estimates in columns (1) and (4); (2) and (5); and (3) and (6), respectively, come from the following model:
$y_{it} = \alpha_0 + \theta_0 Eli_{0it-\tau} + \theta_1 Eli_{1it-\tau} + f_{0g}(z_{it-\tau}; \gamma_{0g}) + \alpha_1 T_{1it-\tau} + f_{1g}(z_{it-\tau}; \gamma_{1g}) \times T_{1it-\tau} + u_{it}$.

Obesity. Table 8 presents the lagged cumulative effects of Medicaid eligibility on the probability of being obese. Columns (1) and (2) show that Medicaid eligibility has a negative impact both on children in the low and in the high income groups after two years of being eligible. However, the effect only persists in the medium run for children of the high income group between 5 and 12 years old. Column (5) indicates that a child eligible for Medicaid in a given period is 13 percentage points more likely to be obese after 5 years than a similar but non-eligible child.

6.3 Channels

The effects of Medicaid on preventive health care utilization show a clear pattern: Medicaid is more likely to increase utilization among children in the low income group but not among the high income group. This differential effect on utilization may be one of the channels through which Medicaid differentially affects children in the low and in the high income groups. In particular, higher utilization of preventive health care may explain why making a child in the low income group eligible for Medicaid makes him more likely to be in excellent health after 3-5 years compared to a similar non-eligible child.

Table 8: Lagged cumulative effects of Medicaid Eligibility on children's health. *Dep. Var: Obesity.*

	Thresholds<185% ($\text{Eli}_t \times 1\{T < 185\}$)			Thresholds[185-250]% ($\text{Eli}_t \times 1\{185 \leq T \leq 250\}$)		
Time Elapsed	5-18 years old (1)	5-11 years old (2)	12-18 years old (3)	5-18 years old (4)	5-11 years old (5)	12-18 years old (6)
1 years (θ_1)	0.061	0.139**	-0.119	0.007	-0.119	-0.021
	(0.056)	(0.067)	(0.110)	(0.048)	(0.110)	(0.071)
2 years (θ_2)	0.119**	0.141**	0.132	0.097*	0.132	0.048
	(0.055)	(0.064)	(0.105)	(0.055)	(0.105)	(0.079)
3 years (θ_3)	-0.03	-0.056	0.030	0.069	0.030	0.012
	(0.051)	(0.064)	(0.085)	(0.056)	(0.085)	(0.084)
4 years (θ_4)	0.020	0.079	-0.106	0.066	0.083	-0.094
	(0.049)	(0.064)	(0.077)	(0.053)	(0.065)	(0.076)
5 years (θ_5)	0.022	0.045	0.012	0.101*	0.130*	-0.027
	(0.042)	(0.050)	(0.079)	(0.056)	(0.073)	(0.086)

Notes: Robust standard errors (in parenthesis) are clustered at the family level. All regressions are linear probability models and all include a polynomial of order 4 of log of income in period $t - \tau$, age, and family size, year and state dummies. The sample is restricted to observations with family income levels that falls within the bandwidth of $\pm$ 30 thousand dollars from the threshold in period $t - \tau$. The intention to treat estimates in columns (1) and (4); (2) and (5); and (3) and (6), respectively, come from the following model:
$y_{it} = \alpha_0 + \theta_0 Eli_{0it-\tau} + \theta_1 Eli_{1it-\tau} + f_{0g}(z_{it-\tau}; \gamma_{0g}) + \alpha_1 T_{1it-\tau} + f_{1g}(z_{it-\tau}; \gamma_{1g}) \times T_{1it-\tau} + u_{it}$.

If Medicaid does not affect the utilization of preventive services of children in the high income group, then the question is why some negative effects on the health of children in this group persist in the medium run, e.g., Medicaid reduces the probability of being in excellent health even after two periods and it increases the probability of being obese even after 5 periods. A second channel consistent with this result is the "quality" channel, that is, differences in the quality of health care the child has access to through Medicaid relative to the counterfactual situation without Medicaid.

Although PSID and CDS datasets do not provide information about the quality of private insurance to directly test whether the quality channel is operating, there is still indirect evidence consistent with this hypothesis. First, children in higher income families are more likely to have private insurance coverage as shown in the Table 1 of Section 4. Indeed, the data show that the counterfactual situation without Medicaid is different across income groups, and it is more likely that a non-eligible child has private coverage the higher the family income.

Second, the quality of care that families have access to through a private insurance may increase with income, i.e., health insurance quality is a normal good. Then, it is more likely that a non-eligible child has a better quality private coverage the higher the income. This is not directly observable, but an implication is that higher income families should be less likely to enroll their children in Medicaid when they are eligible compared to lower income families, because higher income families have better private options they can pay for. Consistent with this implication, I show in Section 5.2 that the higher the eligibility threshold (i.e., the higher the family income) the lower the "jump" in the probability of participation in Medicaid despite being eligible. Another implication of the quality of the private insurance being a normal good is that the difference in the quality of health care obtained through a private insurance and through Medicaid should be increasing in income. Therefore, if a high income family is induced to drop a private insurance in favor of Medicaid, this may imply a drop in health care quality and may have a negative impact on their child's health. My empirical results of the effects of Medicaid on the high income children's health are consistent with this hypothesis.

Finally, there exist some evidence showing that Medicaid may provide lower quality of care than some private insurances. According to the annual State of Health Care Quality Report of the National Committee for Quality Assurance (NCQA), Medicaid plans tend to perform worse, on average, than commercial plans in some iportant quality dimensions, such as whether physicians regularly keep track of children's health by documenting their BMI, or whether during the visit physicians give counseling about nutritions issues and guidance about recommended levels of physical activity to maintain children's health.[23] For instance, during 2010 the percentage of children between 2 to 17 years old who had an outpatient visit with a primary care physician and who had documentation of the BMI percentile, received counseling for nutrition, or received counseling for physical activity were 30.3%, 41.9% and 32.5% for enrolees in Medicaid plans, versus 35.4%, 41.0%, and 36.5% for enrolees in commercials plans (NCQA, 2010). Another measure of queality is whether physicians follow the recommended protocols to treat certain illnesees such as pharyngitis or athsma.[24] According to the NCQA report, the percentage of children between 2 and 18 years old who were diagnosed with pharyngitis and received an appropriate testing was 59.0% in Medicaid versus 74.7% in commercial plans, while the percentage of Medicaid patients with persistent asthma who were prescribed medications acceptable as primary therapy for long-term control of asthma was lower than for patients enrolled in commercial plans (89.6% in Medicaid versus 96.4% in commercial plans, for children between 5 and 9 years old; and 87.0% in Medicaid versus 92.9% in commercial plans, for children between 10 and 17 years old) (NCQA, 2007).

There is some research also providing evidence of a lower quality of Medicaid relative to private insurance. For instance, the amount of time that a doctor spends on average with a Medicaid patient during a visit is lower than for a privately insured patient, as shown by Decker (2007). She finds that in states where Medicaid pays lower fees the amount of time a doctor spends with Medicaid patients is lower relative to privately insured patients. Also in these states physicians are less likely to want to see a Medicaid patient. Hence, a Medicaid beneficiary not only finds more difficult to locate a physician willing to see him, but also the quality of care he receives, measured by the duration of the visit, is also lower than that received by a privately insured patient. Cunningham and O'Malley (2009) also find that Medicaid reimbursement delay affects physicians' willingness to accept Medicaid patients. They show that delays in reimbursement can offset the effects of high Medicaid fees, thereby lowering participation to levels that are closer to those in states with relatively low fees.

[23]The State of Health Care Quality report is produced annually by NCQA to monitor and report on performance trends over time, track variations in patterns of care and provide recommendations for future quality improvement. This report shows indicators coming from The Healthcare Effectiveness Data and Information Set (HEDIS), a tool used by more than 90 percent of America's health plans to measure performance on important dimensions of care and service.

[24]The recommended testing for pharyngitis consist on giving an antibiotic and performing a Group A streptococcus test for the episode.

7 Conclusion

In this paper I analyze the effects of Medicaid on children's health care utilization and health outcomes. I estimate the causal effects of Medicaid taking advantage of Medicaid eligibility rule that generates a discontinuity in the probability of participating in Medicaid. In my analysis I account for potential heterogeneous effects of Medicaid on the health of children with different levels of family income, which is possible due to the variability of eligibility thresholds across states, time, and age groups.

My results highlight the importance of disaggregating the effects of Medicaid depending on the family income level when drawing any conclusions about the effects of the program. Indeed, my findings indicate that Medicaid induces a higher utilization of preventive medical care for the group of children with family income below 185% of the poverty line (the low income group) while it does not produce any significant change for the group of children with family income between 185% and 250% of the poverty line (the high income group).[25] The results also indicate that in the medium run −between 1 and 5 years after being eligible− Medicaid is more likely to have some persistent positive effects on some measures of the low income children's health, while is more likely to have persistent negative effects on the high income children's health.

I proposed two possible channels to explain the differential impact of Medicaid on children's health outcomes in the medium run which are consistent with the findings: the "utilization" channel and the "quality" channel. On the one hand, the utilization channel, according to which Medicaid increases preventive health care utilization and this translates later into better health outcomes, may be the principal mechanism explaining the positive effect on the low income group. On the other hand, the quality channel may be more suitable to explain the negative impact of Medicaid on the high income group. The quality channel explanation states that targeting higher income families with Medicaid may induce a crowding out effect and, although it might not affect health care utilization, it might affect the quality of care a child can have access to. This switch may have undesirable health consequences for children as long as there are health care quality differences between Medicaid and private insurances.[26]

These findings can provide a guide for improving the design and targeting of Medicaid. Medicaid is an effective tool to improve health care access and health outcomes of low income children. However, it can be generating potential conflicts when targeting higher income families. The results of the paper may suggests that the eligibility thresholds are set too high in some states and improvements can be archived reducing them. However, the effects estimated here are performed over a narrow set of health measures and a broader number should be consider to draw this strong policy conclusion. There could be still room for quality improvements, without involving budgetary changes, that may help to reduce the negative unintended effects of Medicaid on higher income children. For example, better monitoring of simple practices that

[25] I cannot draw any conclusions for the group of children with family income higher that 250% the poverty line, because the required condition to apply the fuzzy RD design, i.e., the probability of paticipating in Medicaid as a function of family income changes discontinuously at the threshold, is not satisfied.

[26] Even when Medicaid may also induce some crowding out effect in the low income group, this may not have an unintended effect on children's health. The reason is that if insurance quality is a normal good, then this group is more likely to buy, in the absence of Medicaid, low quality private insurances. Hence, for the low income group switching into Medicaid is more likely to imply an increase in the quality of care they can acquire.

physicians treating Medicaid patients should follow, may lead to better outcomes. Particularly, improving the percentage of physicians that document the BMI and give counseling for nutrition and physical activity may be a cost-effective way to reduce the incidence of obesity on Medicaid eligible children.

References

Almond, D. and J. Currie (2010): "Human Capital Development Before Age Five," NBER Working Papers 15827, National Bureau of Economic Research, Inc.

Black, D., J. Galdo, and J. Smith (2005): "Evaluating the regression discontinuity design using experimental data," *mimeo*, University of Michigan.

Bloom, H. (2009): "Modern Regression Discontinuity Analysis," *mimeo*, MDRC Working Papers on Research Methodology.

Card, D. and L. D. Shore-Sheppard (2004): "Using Discontinuous Eligibility Rules to Identify the Effects of the Federal Medicaid Expansions on Low-Income Children," *The Review of Economics and Statistics*, 86, 752–766.

Carneiro, P. and R. Ginja (2009): "Preventing Behavior Problems in Childhood and Adolescence: Evidence From Head Start," *Mimeo*.

Cellini, S. R., F. Ferreira, and J. Rothstein (2010): "The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design," *The Quarterly Journal of Economics*, 125, 215–261.

Cunningham, P. and A. O'Malley (2009): "Do reimbursement delays discourage Medicaid participation by physicians?" *Health Affairs*, w17–w28.

Currie, J. (2009): "Healthy, Wealthy, and Wise: Socioeconomic Status, Poor Health in Childhood, and Human Capital Development," *Journal of Economic Literature*, 47, 87–122.

Currie, J., S. Decker, and W. Lin (2008): "Has public health insurance for older children reduced disparities in access to care and health outcomes?" *Journal of Health Economics*, 27, 1567 – 1581.

Currie, J. and J. Gruber (1996): "Health Insurance Eligibility, Utilization of Medical Care, and Child Health," *The Quarterly Journal of Economics*, 111, 431–66.

Cutler, D. M. and J. Gruber (1996): "Does Public Insurance Crowd Out Private Insurance?" *The Quarterly Journal of Economics*, 111, 391–430.

Decker, S. (2007): "Medicaid physician fees and the quality of medical care of Medicaid patients in the USA," *Review of Economics of the Household*, 5, 95–112.

Grieger, L., S. Danziger, and R. Schoeni (2009): "Accurately measuring the trend in poverty in the United States using the Panel Study of Income Dynamics," *Journal of Economic and Social Measurement*, 34, 105–117.

Gruber, J. and K. Simon (2007): "Crowd-Out Ten Years Later: Have Recent Public Insurance Expansions Crowded Out Private Health Insurance?" Working Paper 12858, National Bureau of Economic Research.

Hahn, J., P. Todd, and W. Van der Klaauw (2001): "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design," *Econometrica*, 69, 201–09.

Ham, J. C. and L. Shore-Sheppard (2005): "The effect of Medicaid expansions for low-income children on Medicaid participation and private insurance coverage: evidence from the SIPP," *Journal of Public Economics*, 89, 57–83.

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

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

Kaiser Commission on Medicaid and the Uninsured (2010): "Medicaid: A Primer," Tech. rep., Kaiser Family Foundation.

Koch, T. (2010): "Using RD Design to Understand Heterogeneity in Health Insurance Crowd-Out," *Working Paper*.

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

Lo Sasso, A. T. and T. C. Buchmueller (2004): "The effect of the state children's health insurance program on health insurance coverage," *Journal of Health Economics*, 23, 1059–1082.

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

NCQA (2007): "The State of Health Care Quality 2007," Tech. rep., National Committee for Quality Assurance.

——— (2010): "The State of Health Care Quality 2010," Tech. rep., National Committee for Quality Assurance.

van der Klaauw, W. (2002): "Estimating the Effect of Financial Aid Offers on College Enrollment: A Regression-Discontinuity Approach," *International Economic Review*, 43, 1249–1287.

A Manipulation of the Assignment Variable

This appendix discusses a potential threat to the validity of the RD design posed by the possibility that families manipulate their income in order to qualify for Medicaid benefits. I implement the test suggested by McCrary (2008) to test the validity of the RD assumption stating that families do not sort around the eligibility threshold. I perform the test on a sample that restricts to observations in the years 1997, 2002, and 2007, which are the years for which I observe children's outcomes.[27]

Family income is normalized subtracting the corresponding eligibility threshold. The McCrary test follows a two-step procedure: in the first step, the assignment variable −family income− is partitioned into equal spaced bins of width b and the frequencies are computed within those bins. The second step smooths the histogram using local linear regression. The midpoints of the histogram bins are treated as a regressor and the normalized number of observations falling into the bins are treated as a dependent variable in a local linear regression. To accommodate the potential discontinuity in the density, local linear smoothing is conducted separately for the bins to the right and left of the point of potential discontinuity and a triangle kernel is used, with bandwidth h, defining which observations are included in the regression (McCrary, 2008).

The parameter of interest is the log difference in height of the density function, $f(z)$, just below and just above the the threshold, i.e., $\theta = \ln \lim_{z \uparrow z_0} f(z) - \ln \lim_{z \downarrow z_0} f(z) = \ln f^- - \ln f^+$. Under standard nonparametric regularity conditions McCrary (2008) shows that $\hat{\theta} = \ln \hat{f}^- - \ln \hat{f}^+$ is consistent and asymptotically normal.[28] Figure 3 graphically displays the result of the density discontinuity test at the cutoff for different samples. Figure A presents the density estimate for the full sample. The estimate of $\hat{\theta}$ indicates that the log difference in the height of the density function at the threshold is 0.024 (standard error 0.118). The test suggests no discontinuity in the density at the normalized threshold (t-statistic of 0.20).

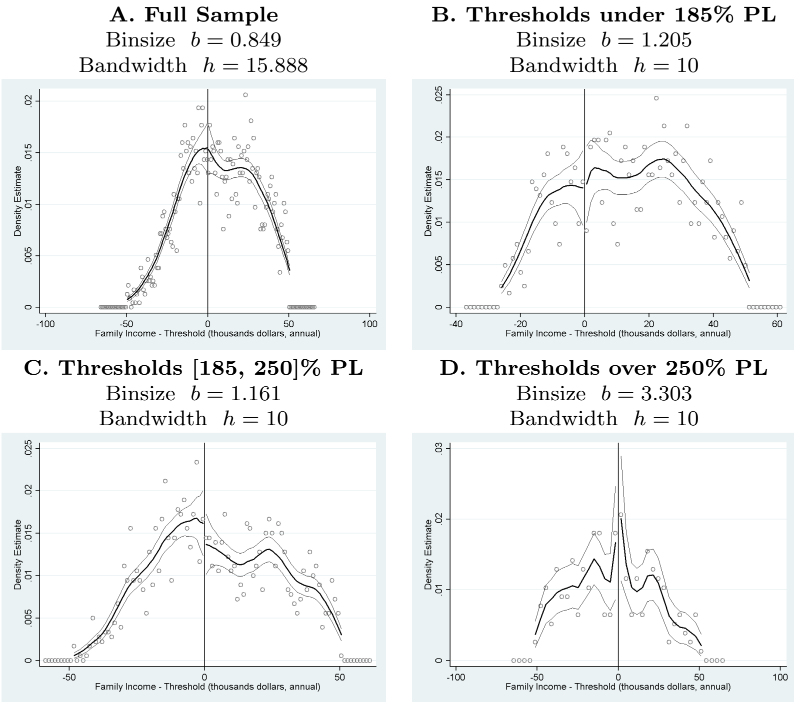
McCrary (2008) indicates that the good performance of $\hat{\theta}$ does not require a careful choice of the binsize, b, in the first stage, by it does require a good choice of the bandwidth, h, in the second stage. I perform the test choosing a variety of bandwidths and keeping the binsize fix. The results of the test for the full sample are reported in Panel A of Table 9. In all cases the hypothesis is not rejected.

To check whether the incentives to manipulate family income vary across different eligibility thresholds I perform the test on three subsamples. The first subsample consists of families who reside in states where the eligibility threshold is lower than 185% the poverty line (Panel B of Figure 3 and Panel B of Table 9), the second subsample considers families who reside in states with eligibility thresholds between 185% and 250% the poverty line (Panel C of Figure 3 and Panel C of Table 9), and finally, a the third subsample consists of families residing in states with thresholds above 250% the poverty line (Panel D of Figure 3 and Panel D of Table 9). Although the graphs show that there could be a greater incentive to manipulate the income

[27] The results are robust when using a sample that considers all years between 1991 and 2007, which is the period for which I can keep track in PSID the family income of the children in my sample. Results available upon request.

[28] I estimate $\hat{\theta}$ using the software `DCdensity.ado` available from McCrary that creates the `DCdensity` command for STATA.

Figure 3: Testing Manipulation of Assignment Variable. Years 1997, 2002, 2007.



Dots are density with the indicated binsize (in thousands dollars). Solid lines are predictions from local linear regressions using triangle kernel with indicated h and b. Standard errors, binsize b, and the bandwidth h are calculated as in McCrary (2008).

when the thresholds are [185, 250]%, in all the cases the test fail to reject the null hypothesis of no discontinuity at the threshold.

B Robustness Analysis of the Discontinuity in the Probability of Participating in Medicaid

In this appendix I perform a robustness analysis to show that the probability of participating in Medicaid as a function of family income is discontinuous at the eligibility threshold. I use the sample that considers family income and Medicaid participation for the whole period 1991 and 2007. Table 10 shows the estimated jump for different parametric specifications, confirming the pattern of Section 5.2.

Given that almost all states have thresholds set below 185% and between 185% and 250% at least once during the period 1991-2007, as it is shown in Table 11, I can extrapolate these results and say that on average children in higer income families are less likely to participate in Medicaid.

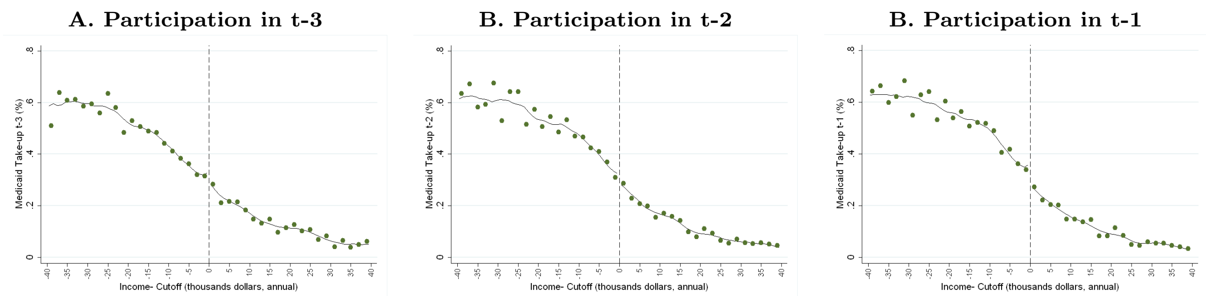
Table 9: McCrary (2008) test for manipulation of assignment variable. Years 1997, 2002, 2007.

		Bandwidth h (thousands)					
	Automatic	10	8	6	4	2	1
A. Full Sample (b=0.849)							
Automatic Bandwidth h	(*15.88*)						
$\hat{\theta}$	0.002	0.020	-0.004	-0.034	-0.031	0.258	0.092
se	0.118	0.150	0.169	0.197	0.250	0.376	0.501
t-statistic	0.020	0.131	-0.024	-0.171	-0.124	0.685	0.184
B. Thresholds [185, 250]% PL(b=1.161)							
Automatic Bandwidth h	(*17.88*)						
$\hat{\theta}$	-0.109	0.010	0.003	0.192	0.734	1.435	0.492
se	0.188	0.264	0.309	0.371	0.512	0.846	0.920
t-statistic	-0.579	0.037	0.009	0.519	1.433	1.696	0.535
C. Thresholds [185, 250]% PL(b=1.161)							
Automatic Bandwidth h	(*21.62*)						
$\hat{\theta}$	0.211	0.157	0.121	0.025	-0.081	0.283	0.143
se	0.137	0.204	0.226	0.261	0.323	0.434	0.633
t-statistic	1.546	0.770	0.535	0.096	-0.251	0.652	0.226
D. Thresholds over 250% PL(b=3.303)							
Automatic Bandwidth h	(*27.63*)						
$\hat{\theta}$	-0.194	-0.145	-0.053	-0.053	-0.134	-0.134	-
se	0.324	0.427	0.457	0.528	0.729	1.030	-
t-statistic	-0.600	-0.339	-0.115	-0.100	-0.183	-0.130	-

Notes: $\hat{\theta} = \ln \hat{f}^{-} - \ln \hat{f}^{+}$ estimates the discontinuity in the density function of the assignment variable at the threshold. A positive and statistically significant value of $\hat{\theta}$ may be an indicator of sorting around the threshold. "Automatic" refers to the bandwidth obtained using the automatic selection procedure proposed by McCrary (2008).

B.1 Placebo test for discontinuity

Figure 4: Discontinuity in the probability of participation. Placebo tests.



C Sensitivity Analysis: Model Specification

In this section I present alternative specifications to those obtained in Section 6. Here I consider different orders of polynomials of the income function to check the robustness of the results for utilizations. Robustness analysis for health measures are available upon request.

Table 10: Participation Equation. "Jump" at the threshold. Period 1991-2007.

	Bandwidth (thousands dollars)				
	±50	±30	±20	±15	±2
A. Full sample					
Polynomial Order					
One	0.208***	0.199***	0.158***	0.093***	0.074***
	(0.015)	(0.015)	(0.016)	(0.015)	(0.028)
Two	0.126***	0.104***	0.083***	0.063***	0.075***
	(0.014)	(0.014)	(0.015)	(0.015)	(0.028)
Three	0.121***	0.086***	0.063***	0.057***	0.070***
	(0.014)	(0.014)	(0.015)	(0.015v	(0.028)
Four	0.112***	0.086***	0.064***	0.056***	0.073***
	(0.014)	(0.014)	(0.015)	(0.015)	(0.028)
B. Model Interacted					
Polynomial Order					
One					
$Eli_t \times 1\{T < 185\}$	0.279***	0.275***	0.235***	0.151***	0.117***
	(0.021)	(0.021)	(0.022)	(0.023)	(0.042)
$Eli_t \times 1\{185 \leq T \leq 250\}$	0.188***	0.178***	0.106***	0.095***	0.059
	(0.022)	(0.023)	(0.022)	(0.024)	(0.041)
$Eli_t \times 1\{T > 250\}$	0.022	0.051	0.029	0.053	-0.048
	(0.037)	(0.035)	(0.040)	(0.047)	(0.081)
Two					
$Eli_t \times 1\{T < 185\}$	0.178***	0.157***	0.141***	0.113***	0.114***
	(0.021)	(0.022)	(0.023)	(0.024)	(0.042)
$Eli_t \times 1\{185 \leq T \leq 250\}$	0.115***	0.098***	0.073***	0.073***	0.056
	(0.020)	(0.021)	(0.023)	(0.023)	(0.042)
$Eli_t \times 1\{T > 250\}$	0.036	0.041	0.013	0.039	-0.054
	(0.039)	(0.039)	(0.039)	(0.046)	(0.081)
Three					
$Eli_t \times 1\{T < 185\}$	0.174***	0.128***	0.105***	0.096***	0.121***
	(0.020)	(0.022)	(0.024)	(0.025)	(0.042)
$Eli_t \times 1\{185 \leq T \leq 250\}$	0.112***	0.076***	0.068***	0.063***	0.056
	(0.020)	(0.022)	(0.023)	(0.023)	(0.042)
$Eli_t \times 1\{T > 250\}$	0.053	0.031	0.017	0.037	-0.070
	(0.038)	(0.038)	(0.040)	(0.046)	(0.089)
Four					
$Eli_t \times 1\{185 \leq T \leq 250\}$	0.163***	0.130***	0.107***	0.100***	0.129***
	(0.021)	(0.022)	(0.024)	(0.025)	(0.042)
$Eli_t \times 1\{T > 250\}$	0.109***	0.087***	0.068***	0.064***	0.055
	(0.020)	(0.022)	(0.023)	(0.024)	(0.042)
$Eli_t \times 1\{T < 185\}$	0.058	0.036	0.017	0.039	-0.069
	(0.039)	(0.038)	(0.041)	(0.046)	(0.090)
N	22,701	17,857	13,391	10,411	1,426

Notes: **Panel A:** Each entry comes from a separate linear probability model $M_{i,t} = \pi_0 + \pi_1 Eli_{it} + k_{1g}(z_{it};\alpha_{1g}) + u_{it}$. All regressions include a polynomial of the indicated order of log income, age, and family size; year and state dummies. **Panel B:** Each entry comes from a separate linear probability model $M_{it} = \gamma_0 + \sum_{j=1}^{2} \gamma_j T_{j,it} + \sum_{j=0}^{2} \pi_j Eli_{j,it} + k_{0g}(z_{it};\alpha_{0g}) + \sum_{j=1}^{2} k_{jg}(z_{it};\alpha_{jg}) \times T_{j,it} + u_{it}$. All regressions include a polynomial of the indicated order of the log income, age, and family size; year and state dummies. Robust standard errors (in parenthesis) are clustered at the family level. In each column the sample is restricted to observations with family income levels that falls within the bandwidth indicated.

Table 11: Eligibility Thresholds by State

State	The state has at least once, during the period 1991-2007, a threshold: under 185 % the FPL	[185,250] % the FPL	over 250 % the FPL
Alabama	Yes	Yes	No
Alaska	Yes	Yes	No
Arizona	Yes	Yes	No
Arkansas	Yes	Yes	No
California	Yes	Yes	Yes
Colorado	Yes	Yes	No
Connecticut	Yes	Yes	Yes
Delaware	Yes	Yes	No
District of Columbia	Yes	Yes	No
Florida	Yes	Yes	No
Georgia	Yes	Yes	No
Hawaii	Yes	Yes	Yes
Idaho	Yes	Yes	No
Illinois	Yes	Yes	No
Indiana	Yes	Yes	No
Iowa	Yes	Yes	No
Kansas	Yes	Yes	No
Kentucky	Yes	Yes	No
Louisiana	Yes	Yes	No
Maine	Yes	Yes	No
Maryland	Yes	Yes	Yes
Massachusetts	Yes	Yes	No
Michigan	Yes	Yes	No
Minnesota	Yes	Yes	Yes
Mississippi	Yes	Yes	No
Missouri	Yes	Yes	Yes
Montana	Yes	No	No
Nebraska	Yes	Yes	No
Nevada	Yes	Yes	No
New Hampshire	Yes	Yes	Yes
New Jersey	Yes	Yes	Yes
New Mexico	Yes	Yes	No
New York	Yes	Yes	No
North Carolina	Yes	Yes	No
North Dakota	Yes	No	No
Ohio	Yes	Yes	No
Oklahoma	Yes	Yes	No
Oregon	Yes	Yes	No
Pennsylvania	Yes	Yes	No
Rhode Island	Yes	Yes	No
South Carolina	Yes	Yes	No
South Dakota	Yes	Yes	No
Tennessee	Yes	Yes	Yes
Texas	Yes	Yes	No
Utah	Yes	Yes	No
Vermont	Yes	Yes	Yes
Virginia	Yes	Yes	No
Washington	Yes	Yes	No
West Virginia	Yes	Yes	No
Wisconsin	Yes	Yes	No
Wyoming	Yes	Yes	No
	51	49	10

Table 12: Contemporaneous effects of Medicaid on utilization. Sensitivity analysis for different model specifications and window widths. *Dep. Var.: The child has visited a doctor for a routine health check-up in the last 12 months.* Children between 5 and 18 years old. Years 1997, 2002, and 2007.

	Bandwidth (thousands dollars)			
Polynomial Order	±30	±20	±15	±2
A. Intention to treat				
One				
$Eli_t \times 1\{T < 185\}$	0.147***	0.159***	0.174***	0.044
	0.051	0.058	0.063	0.138
$Eli_t \times 1\{185 \leq T \leq 250\}$	0.043	-0.010	-0.048	-0.093
	0.042	0.047	0.054	0.114
Two				
$Eli_t \times 1\{T < 185\}$	0.131**	0.177**	0.183**	0.044
	0.061	0.069	0.073	0.138
$Eli_t \times 1\{185 \leq T \leq 250\}$	-0.013	-0.010	-0.008	-0.138
	0.045	0.060	0.067	0.126
Three				
$Eli_t \times 1\{T < 185\}$	0.132**	0.174**	0.175**	0.045
	0.063	0.070	0.075	0.136
$Eli_t \times 1\{185 \leq T \leq 250\}$	0.000	-0.009	-0.020	-0.108
	0.050	0.059	0.065	0.127
Four				
$Eli_t \times 1\{T < 185\}$	0.138**	0.157**	0.169**	0.042
	(0.065)	(0.072)	(0.076)	(0.136)
$Eli_t \times 1\{185 \leq T \leq 250\}$	-0.003	-0.005	-0.022	-0.118
	(0.050)	(0.060)	(0.065)	(0.131)
B. Outcome Equation. IV-RD				
One				
$M_t \times 1\{T < 185\}$	0.396***	0.446**	0.553**	0.183
	0.137	0.202	0.261	0.402
$M_t \times 1\{185 \leq T \leq 250\}$	0.159	-0.108	-0.136	-0.761
	0.163	0.276	0.260	0.634
Two				
$M_t \times 1\{T < 185\}$	0.373	0.506**	0.618**	0.193
	0.229	0.234	0.303	0.390
$M_t \times 1\{185 \leq T \leq 250\}$	-0.079	-0.011	0.026	-0.870
	0.359	0.273	0.300	0.715
Three				
$M_t \times 1\{T < 185\}$	0.415*	0.505**	0.581*	0.161
	0.222	0.240	0.309	0.380
$M_t \times 1\{185 \leq T \leq 250\}$	0.088	0.024	-0.013	-0.698
	0.259	0.258	0.294	0.627
Four				
$M_t \times 1\{T < 185\}$	0.455*	0.524**	0.517*	-0.193
	0.233	0.225	0.300	0.403
$M_t \times 1\{185 \leq T \leq 250\}$	0.080	0.082	-0.071	-0.985
	0.263	0.237	0.291	1.344
N	1992	1441	1102	156

Notes: Robust standard errors (in parenthesis) are clustered at the family level. All regressions include a polynomial of the indicated order of log income, age, and family size, and year and state dummies. In each column the sample is restricted to observations within the bandwidth indicated. The intention to treat estimates in each column come from the following model: $y_{it} = \alpha_0 + \theta_0 Eli_{it} + \theta_1 Eli_{it} \times T_{1it} + f_g(z_{it}; \gamma_g) + \alpha_1 T_{1it} + f_g(z_{it}; \gamma_{1g}) \times T_{1it} + u_{it}$. The IV-RD estimates in each column come from the following model:
$y_{it} = \alpha_0 + \beta_0 M_{it} + \beta_1 M_{it} \times T_{1it} + k_{2g}(z_{it}; \alpha_{1g}) + \alpha_1 T_{1it} + k_{2g}(z_{it}; \alpha_{1g}) \times T_{1it} + u_{it}$, where eligibility instruments for Medicaid coverage.