**Joint Center for Housing Studies**
**Harvard University**

# Housing Assistance and Student Achievement in Low-Income Households in Chile

**Felipe Kast**

**December 2009**
**W09-8**

*Felipe Kast is the 2007 recipient of the John R. Meyer Dissertation Fellowship*

The author thanks Alberto Abadie, Daniel Hojman, Sandy Jencks, Dan Levy, Jeffrey Liebman, Slaven Razmilic, Nicolas Retsinas, Tristan Zajonc and Richard Zeckhauser for helpful comments and discussions. I am also is grateful to the Harvard Joint Center for Housing Studies and the Harvard Taubman Center for State and Local Government for financial support, to the Ministry of Public Housing of Chile, to the Observatorio Habitacional for efficient provision of information, and especially to Michele Favre for excellent research assistance..

**Abstract**

This paper examines the impact of publicly provided housing unit on student achievement in the context of low income households of Chile. It differs from previous studies evaluating externalities of promoting homeownership by using a regression discontinuity approach in which the underlying assumptions required for a proper identification can be tested. Data taken three to six years after the provision of housing units suggest that the treatment increased by 0.39 years the level of education for the population under 25 years old. Among the mechanisms that could explain this phenomenon I find that three channels are triggered by the treatment. There is a cash transfer equivalent to the market value of the housing unit, there is a positive impact on the housing conditions of the beneficiaries, and the probability of being homeowner is 20% higher.

## Introduction

Promoting homeownership is an old and popular tradition in social policy. From Franklin D. Roosevelt's belief that "[a] nation of homeowners is unconquerable" to President Bush's statement that "[o]wning a home lies at the heart of the American dream," there has been a constant enthusiasm for this policy. Far from being merely rhetoric, a preference for homeownership is reflected in U.S. policy: the I.R.S. allows home owners to deduct mortgage interest from their taxable income, the federal tax system provides a significant subsidy to owner-occupied housing, and government-sponsored enterprises — Fannie Mae, Freddie Mac and Federal Home Loan Banks — provide cheap financing for home buyers. These policies give renters an incentive to become homeowners earlier than otherwise feasible within their economic status. This, in turn, may affect their financial stability, labor supply and residential mobility.

The critical role played by housing policy in the recent global financial crisis has raised many questions about the social value of subsidizing homeownership, some of which this article will examine in the context of public provision of housing units in developing countries. The rationale of these policies can be traced back to the end of the nineteen century when government councils of overcrowded cities provided housing assistance with the intention of preventing epidemics and health hazards. Nowadays, however, the justification for housing assistance is not necessarily to improve health but to promote a number of potential positive externalities. Rosen (1985) questioned the validity of the externality argument and a broad range of empirical evaluations estimating these externalities emerged. In general these studies find positive externalities, but a comprehensive review of this literature by Dietz and Haurin (2003) reveals that most studies lack a clear identification strategy and usually contain serious selection problems. Rosen's original question has returned full-face as this past year economist Paul Krugman asked, "why should ever-increasing homeownership be a policy goal?"

An inclination in favor of subsidizing homeownership is not an exclusive feature of developed countries. Fisher and Jaffe (2003) document that homeownership rates are usually higher in low income countries, while Buckley and Kalarickal (2006) show that housing subsidies as a portion of GDP are also higher in low income countries. Although the theoretical reason for favoring housing policies in developing countries is similar to that of developed countries – the possibility of generating positive externalities – in practice the estimated parameters may be quite different if initial housing conditions are more extreme. Rosen itself had argued that “if externalities were important anywhere, it would be in the slums, where housing density is very high.”

In this study I estimate the impact of publicly provided new housing units on education by taking advantage of a targeting system for housing vouchers implemented in Chile in the early eighties. The identification strategy flows directly from the design of the Chilean housing subsidy. The subsidy consists of a voucher that covers between 60% and 95% of the total cost of a new housing unit and can be used in a special market where private firms can offer new units if they comply with the required specifications. The program is means-tested, and during the application process a poverty score card is used for targeting purposes. Applicants do not know in advance the poverty cutoff point, and indicators are obtained from different dimensions (poverty intensity, waiting time, housing needs and savings). Finally, based on budget availability, the Ministry of Housing decides how many vouchers are going to be delivered in a particular application process.

In this particular setting the use of the housing vouchers associated may be associated with four components: (1) income effect; the market value of the voucher is about 1.5 to 3 times the annual income of a beneficiary, (2) housing effect; new housing units may have better conditions, (3) neighborhood effect; new housing units are located in a different location, normally where the price of the land is cheaper, (4) homeownership effect; the voucher may induce early ownership of a housing unit.

The impact of the treatment is estimated by comparing the outcome for applicants with scores just above and below the threshold generated in each assignment process. This corresponds to the “Intent-to-Treat” (ITT) effect, and from a policy point of view this parameter tells us the impact of offering a housing voucher in the margin. Since I observe partial compliance with the treatment I am also interested in estimating the effect of actually using the housing voucher, which is known as the effect of the “Treatment-on-Treated” (TOT). Here I don´t have access to data at the individual level on who actually used the vouchers, but I can recover a proxy of this parameter using aggregate take up rates from administrative data provided by the Ministry of Housing.

Concretely, I am interested in estimating the impact of publicly provided housing unit on the level of education obtained by the population under 25 years old. After three to six years I find that the treatment increased by 0.39 years the level of education obtained by the population under 25 years old. Among the mechanisms that could explain this phenomenon the evidence shows that housing conditions are improved, the probability of being homeowner goes up, and the characteristics of the neighborhood aren´t affected by the treatment. Thus, the impact of the treatment on education is justified by the combination of the income, housing and homeownership effect

These results are consistent with the positive income effect on education found by Currie and Yelowitz (2000) and also provide valuable information about a relevant channel that may affect educational outcomes in developing countries. While these results are also consistent with current evidence in developed countries suggesting that promoting homeownership has a positive impact on education by reducing residential mobility (Aaronson, 2000) , or by generating social capital (Glaeser and DiPasquale, 1998), the magnitude of the effect suggests that in developing countries education can be directly constrained by precarious housing conditions and income needs.

The remainder of the paper is as follows. I first provide a review of the literature in Section 2. Then I provide background information about the treatment and the assignment process in Section 3, and the empirical strategy is described in detail in Section 4. A description of the data used in the analysis can be found in Section 5, the results are presented in Section 6, and Section 7 concludes.

## Literature Review

In this section I discuss previous evidence about the impact of these components (income, housing, neighborhood and homeownership) on educational outcomes.

On neighborhood effects the most well-known large scale study is the randomized housing mobility program "Moving to Opportunity" (MTO) by Katz, Kling and Liebman (2007). Data four to seven years after the provision of the vouchers show no significant effect on adult earnings, employment and physical health. However, the authors do find a positive impact on mental health for adults and for younger women. A positive effect on the education and physical health of young women was also found, but this effect was counterbalanced by a negative impact on contemporary young men.

A similar result was obtained from an evaluation of long-run neighborhood effects by Oreopoulos (2003). The author finds that neighborhood quality plays little role in determining a youth's future earnings. For the identification strategy, he uses the exogenous allocation of public projects into different types of neighborhoods[1] by matching project addresses with administrative data to track children when they are older than 30 years old.

Another well-identified study that combines neighborhood and housing components was done by Jacob (2005). In this case, he uses an exogenous variation in neighborhood generated by high-rise demolitions in Chicago to examine the impact of high-rise public housing on educational outcomes. He compares the children from these buildings to students from other similar projects and finds no effect on student outcomes. The author claims that because the new neighborhoods resemble the original dwellings, this can be interpreted as the independent effect of high-rise public buildings.

In sum, well identified evidence of neighborhood effects suggests that having a better environment may have an impact on psychological and health outcomes for some subgroups in the population, but outcomes such as education and earnings are not found to be significantly influenced by housing policies of this type.

On the homeownership component  Dietz and Haurin (2003) find that, with a few exceptions,  most of the past 30 year of literature on homeownership effects contains selection problems.  Here I describe the cases where proper identification is plausible.

Glaeser and DiPasquale (1998) find that homeownership has a positive externality on the formation of social capital, using instrumental variables to satisfy internal validity. The main idea is that homeowners invest more in their community relations because the quality of their communities is correlated with the value of their homes. As an instrument they use the average homeownership rate of the individual's income quartile for each race in each state. The problem is that other variables not considered in the model may affect homeownership and social capital concurrently, which would bias the estimated parameters.

Closer to the focus of this study, though not using exogenous variation as identification strategy, Green and White (1997) evaluate the impact of homeownership on children's outcomes. They find that for a 17-year-old teenager from a low-income household owning a house, the

[1] Excerpt describing the variation: "Some projects consist only of high-rise apartments; others are only townhouses. Some accommodate more than 10,000 individuals; others provide shelter to less than 100 individuals. And some projects are located in central downtown, while others are in middle-income areas in the suburbs."

predicted probability of staying in school is 0.86 compared to 0.67 for a similar household renting a unit. They also find that when income increases the difference between the probabilities decreases.  The authors deal with selection by estimating a bivariate profit for an endogenous switching model that explains both homeownership and stay-in-school decision. This parametric self-selection correction, however, imposes some assumptions about the selection equation that are difficult to test.

Aaronson (2000) builds upon the work of Green and White (1997) and focuses on the selection problem using instrumental variables. For homeownership he uses the same instrument proposed by Glaeser and DiPasqualle (1997) and finds that the impact is positive but smaller than the one proposed by Green and White (1997). Furthermore, he introduces residential mobility as an independent variable into the equation and finds that most of the effect comes from this channel. The problem, again, is the validity of the instrument. Aaronson uses residential mobility rates prior to the child turning five as an instrument for mobility rates between age 7 and 16, but the instrument is questionable as long as early mobility is correlated with children's educational outcomes.

Haurin, Parcel and Haurin (2002) uses the relative price of homeownership as an instrument and find that children from homeowners perform better in math and reading exams. They argue that variation in house prices and down payment have an impact on homeownership but not in child outcomes. This is not going to be the case if the variation in house prices has a direct income effect for the household, which in turn may affect educational outcomes.  Instead, we should interpret the estimate as the combined effect of the income and the homeownership effect.

Currie and Yelowitz (2000) are able to isolate the income effect by looking a case where housing assistance takes the form of rent subsidy, and they find a positive impact on education and housing conditions. The authors use children gender composition as the instrument. A family with two boys is assigned a two bedroom apartment, but a family with a boy and a girl gets a three bedroom apartment. Therefore, they argue, the second family should be more likely to apply for the project. A minor concern with the instrument is that gender composition could be correlated not only with the probability of getting housing assistance but also with children's educational outcomes.

The evidence presented above suggests that neighborhood effects do not play a crucial role on education, homeownership effects are not well identified, and income effects are found to have a positive effect on education.

This paper differs from previous evidence in three dimensions. First, by using a regression discontinuity approach I am able to test the underlying assumptions required for a proper identification. Second, the nature of the data provides a unique opportunity to evaluate the impact of the treatment on the different channels that may explain the overall effect on education. Finally, the parameters are estimated in a context of low income families from a developing country.

## Background

Chile introduced a voucher housing policy by the late 1970s[2], and identification emerges directly from the assignment of this subsidy. The policy consists of an up-front capital subsidy that can be used in the private sector covering from 60% and 95% of the cost of a new housing unit. At the beginning, vouchers were granted directly to those living in marginal conditions or slums, but in 1984, the Ministry of Housing implemented an innovative selection system to scale-up the program. The importance of housing public assistance grew steadily over time and according to Pardo (1999), by 1998, about 57% of the total units constructed in Chile (137.043 housing units) received a direct government subsidy.

In the application process, each individual obtains points from four factors: socioeconomic situation, family size, savings and time on waiting list. Applicants are given priority based on their weighted score, each program has different requirements, and some programs are not restricted to a certain socioeconomic group. The amount of the subsidy is fixed, not dependent on the family size or geographic location.

Every year, the Minister of Housing offers different packages with a number of vouchers available, and anyone registered with a valid application can choose to participate. Applicants are sorted by their total score and vouchers are assigned according to their scores. The last

[2] According to Gilbert (2004) Chile was the first country in the world to introduce a housing voucher.

person to get a subsidy generates the cut off that I will use as the source of exogenous variation needed for identification.

Once an applicant gets the voucher, the individual can use or discard it, and it is not unusual to observe the later. There are two reasons. First, the voucher may be insufficient to cover the cost of a housing unit in the preferred location. Second, Chileans can apply to the vouchers as many times as they want, but once the voucher is used they are prevented from applying again. Thus, applicants will discard a voucher if they expect to get a better deal in the future, in terms of better location or a better housing unit.

In this study, I focus my attention on two particular programs: (1) Vivienda Progresiva, created in 1990 as an effort to focus the attention on the poorer population and reduce squatter settlements, and (2) Vivienda Básica, created during the previous decade as a more definitive solution, and then a similar program was added in the mid nineties as a voucher that can be used in the private sector (Nueva Básica). These programs provide a housing unit in a site of 100 sq. meters with access to water, electricity and a sewer system. While Vivienda Progresiva consists of an expandable construction of 23 sq. meters, Vivienda Básica usually provides units of about 40 sq. meters. The monetary value of the voucher accounts for about 1.5 to 3 times the annual income of an eligible family. As part of the requirements applicants are asked to show a valid poverty score card and a minimum amount of savings. For completeness, a detailed description of a typical housing unit for each program is provided in Appendix A.1 and A.2

The government of Chile played an important role promoting the housing market among low income families, but this intervention also affected the land market and the shape of cities. As a way to reduce the cost of construction for low income families the land market was liberalized and many borders of the city eliminated. This policy had the benefit of relaxing budget constraints, but it also generated new challenges in terms of infrastructure and transportation. Most new units were constructed in the borders of the city, and it is not clear whether these neighborhoods concentrate more poverty or delinquency. Here I am able to test whether this assumption about the quality of the new localities is compatible with the data.

The two programs described before were the most important programs targeted at the low-income population between the mid-1980s and the end of the 1990s. A detailed description of the magnitude and evolution of each program can be found in Appendix A.3. By the year 2000, the importance of these two programs began to decrease, so that by 2005, they had been completely replaced. In this study, data was accessed for applicants of the two programs between 1998 and 2003 for the three largest regions of the country: Región Metropolitana, V and VIII región.

**Identification Strategy**

Does public provision of a housing unit have an impact on children's educational outcomes? If housing units were randomly assigned among the applicants this question would be answered by comparing educational outcomes of teenagers living in these houses with those who did not get the treatment. The reality is that houses are not assigned by lottery, and those families who own a house are generally better off. As a result, conducting a simple comparison would generate biased estimates.

The nature of the assignment of vouchers described above provides a clean opportunity to estimate the causal effect of publicly provided housing unit on education level using a Regression Discontinuity design. The intuition is that although housing subsidies are not assigned randomly, the systematic use of a score in the assignment process and the generation of a cutoff mimics random assignment around the threshold. Therefore, comparing the outcome for those just above and below the cutoff provides an unbiased local estimate of the parameter of interest.

The use of the Regression Discontinuity approach is going to be appropriate as long as the unobservable factors are continuously distributed over the score variable, which is likely to occur if agents are not able to manipulate the probability of getting the voucher in a deterministic way. I will formally address this point at the moment of estimating the parameters.

In this setting, comparing the outcome for those just above and below the cutoff corresponds to the "Intent-to-Treat" (ITT) effect, and from a policy point of view this parameter tells us the marginal effect of increasing the number of vouchers offered by the Ministry of Housing. But since compliance with the treatment is partial, and because from a theoretical perspective we are also interested in estimating the effect of actually getting a new housing unit, it is crucial to get an estimate of "Treatment-on-Treated" (TOT) effect. And although I don't have access to individual data that would allow me to identify those who actually used the vouchers, under reasonable assumptions I am able to estimate this parameter using official aggregate take-up information provided by the Ministry of Housing of Chile.

I formally describe the empirical strategy introducing the basics of the Rubin Causal Model [Rubin (1974), Holland (1986)]. I want to estimate the impact of the treatment $D$, public provision of homeownership in this case, on a subsequent outcome $Y$. Under this model $Y_{i0}$ denotes the outcome level without the treatment, and $Y_{i1}$ describes the outcome with the treatment. Let $D_i \in (0,1)$ be the treatment received, with $D_i=1$ if the individual is treated, and $D_i=0$ otherwise. For each individual a vector is observed of covariates described by $(X_i, Z_i)$ , where the probability of assignment to the treatment is at least partially determined by whether $X_i$ , the running variable that in this case is the score obtained in the application, is above or below a fixed threshold $c$. The problem is that it is not possible to observe both parameters for each individual. Instead we observe

$$(1) \quad Y_i = D_i * Y_{i1} + (1 \quad D_i) * Y_{i0}$$

Given the nature of the assignment process, and in order to estimate the "Intent-to-Treat" effect of the policy we can use what is known as the Sharp Regression Discontinuity (SRD). For the use of the SRD the probability of getting the treatment must be fully determined by whether $X$ is above or below $c$. Thus, I define the treatment as "being offered the voucher" ($D_{ITT}$) in order to comply with this condition. For the case of the "Treatment-on-Treated" effect the treatment is defined as "using the voucher," ($D_{TOT}$) and in this case, because a partial compliance is observed, the adequate estimation strategy is known as Fuzzy Regression Discontinuity (FRD) design, and the assignment of the voucher is used as an instrument for the treatment. Similar to the case of instrumental variables, in the FRD design the estimated parameter of interest is valid only for compliers, or those who use the voucher if they get the subsidy, but not otherwise.

For both cases the average causal effect of the treatment is estimated by computing the ratio between the change in outcome and jump in the probability of getting the treatment at the threshold. Equation (2) generalizes these situations using only observations that are closer than $h$ from $c$.

$$(2) \quad \alpha = (E[Y|X \geq c] - E[Y|X \leq c])/(E[D|X \geq c] - E[D|X \leq c]) \mid 1(c - h < X < c + h)$$

For the "Intent-to-Treat" effect the discrete change in the probability of getting the treatment is one, corresponding to a SRD design. Thus, $E[D_{ITT}|X \geq c] - E[D_{ITT}|X \leq c] = 1$ and I am able to estimate αITT according to equation (3).

(3) $Y|X \leq c] \Big|_{c-h<X<c+h})$

For the "Treatment-on-Treated" the discrete change in the probability is less than one, and the proper strategy is a FRD design. Here we can see the indicator of being above the cutoff works as an instrument for the treatment[3], and I can to recover $\alpha_{TOT}$ from equation (4). Here, as it is in the instrumental variables setting, the monotinicity assumption requires that the instrument to be valid must affect participation only in one direction.

$$(4) \quad \alpha_{TOT} = \frac{\alpha_{ITT}}{E[D_{TOT}|X \geq c] - E[D_{TOT}|X \leq c]} \Bigg|_{c-h<X<c+h}$$

For the estimation of $\alpha_{ITT}$ and $\alpha_{TOT}$ I use both a nonparametric and a parametric procedure as a way to evaluate the stability of the findings across different specifications. In principle, the nonparametric approach has the advantage of using only variables located close to the boundary, but using a finite sample it is not possible to know which one of the two approaches has a smaller bias. Lee and Lemieux (2009) argue that we should see these two alternatives as complements rather than substitutes, and that only results that are relatively stable across different specifications should be trusted.

In the nonparametric case I follow the convention of using local linear regressions to estimate the coefficients at both sides of the cutoff point. In practice, and because rectangular kernel is a convenient choice, using local linear regressions corresponds to fitting a linear regression using observations within a distance *h* on either side of the threshold. For computational reasons it is convenient to subtract the cutoff value from the running variable. By doing this the intercepts from the regression on either side provides the value of the regression at the point of interest.

[3] The connection between this strategy and the instrumental variables approach was made explicit by Hahn, Todd and Klaauw (2001)

As it is described above, in this setting each public call made by the Ministry of Housing generates its own cutoff. Therefore each observation in the sample is also indexed by j, the corresponding public call. This means that the estimated parameter $\alpha_{ITT}$ is going to be the weighted average of the parameters estimated from all public calls, $\alpha_{ITT(j)}$.

Equation (3) is therefore computed by running two regressions. On the left hand side of the threshold I run

$$(5) \qquad Y_{ij} = \mu_{lj} + \beta_{lj} * (X_{ij} - c_j) + \varepsilon_{ij} \qquad if\ (c_j - h < X_{ij} < c_j)$$

On the right hand side I run

$$(6) \qquad Y_{ij} = \mu_{rj} + \beta_{rj} * (X_{ij} - c_j) + \varsigma_{ij} \qquad if\ (c_j < X_{ij} < c_j + h)$$

From (5) and (6) I can recover $\mu_{rj}$- $\mu_{lj}$ , which is an unbiased estimator for $\alpha_1(ITT(i)) = E[Y|X \geq c] - E[Y|X \leq c] \mid_1 (c - h < X < c + h)$ . For computation reasons it is convenient to get this parameter running one regression where we pool the data from both sides of the cutoff. That way is possible to estimate the standard errors directly from the regression. Let us then define $T_{ij}$ as a dummy taking value equal to one if observation *"ij"* is above $c_j$ , and zero otherwise. In this case $\pi_j = \mu_{rj} - \mu_{lj}$ from equation (7) is equivalent to $\alpha_{ITT(j)}$.

$$(7) \qquad Y_{ij} = \mu_{lj} + \pi_j * T_{ij} + \beta_{lj} * (X_{ij} - c_j) + (\beta_{rj} - \beta_{lj}) * T_{ij} * (X_{ij} - c_j) + \zeta_{ij} \qquad if\ (c_j < X_{ij} < c_j + h)$$

A relevant feature of this specification is the fact that the slope of the regression can differ on both sides of the cutoff. This is important if we want to avoid data from the left affecting our estimate of the intercept on the right side of the cutoff, and vice versa.

Now I need to go from equation (3) to equation (4), or in other words, from ITT to TOT. In order to do this I need an estimate for the denominator of equation (4). Normally one would obtain this estimate by running a regression similar to equation (7), but replacing the outcome for the treatment variable $D_{TOT}$ on the left hand side of the equation. I do not have access to individual data of those who actually used the vouchers, and I am forced to estimate this parameter using official information provided by the ministry of housing regarding take-up rates for the programs under study. Thus, because $E[D_{TOT}|X \geq c] - E[D_{TOT}|X \leq c]\big|_{c-h<X<c+h}$ is obtained with low precision we should be cautious at interpreting TOT results.

Finally, I need to go from the estimate for a particular public call $\alpha_{ITT(j)}$ to $\alpha_{ITT}$, a summary index equal to the weighted average of the parameters estimated for each public call.

One of the key issues for the nonparametric approach is the definition of the optimal bandwidth (h*). It is always recommended to present the estimates using different bandwidths, but choosing the right window of width around the threshold is probably the most crucial component of the RD estimation strategy. Using large bandwidth reduces the variance but increases the bias, and this bias is going to be significant if the underlying conditional expectation of the outcome is not linear in the range considered by the window h.

Here I identify the optimal bandwidth using the standard "leave one out" cross-validation procedure recommended by Imbens and Lemieux (2008). The procedure goes as follows: take a particular value of h´, and for each observation "*i*" sitting at the left side of the threshold run a local linear regression leaving this observation out, and using its score $X_i$ as the relevant cutoff $(X_i - h < X < X_i)$. Take the estimates of this regression to get the predicted value of the outcome at $X = X_i$ and compute the difference between the predicted value and $Y_i$, the value of the outcome for observation "*i*". Repeat the process for observations at the right hand of the threshold, running a local linear regression over $(X_i < X < X_i + h)$ and finally compute the mean square of the difference between the predicted value and $Y_i$ for every "*i*". The optimal choice is defined by the bandwidth that minimized the mean square.

In the parametric case I run polynomial regressions at each side of the boundary. Instead of selecting one particular functional form I report the coefficients for the second, third and fourth order polynomial regressions. As with the local linear regression, here it is also convenient to run a pooled regression as a way to recover the robust standard errors. For the case of the second order polynomial regression, for example, I am able to recover $\alpha_{ITT(j)}$ from $\pi_j$ running equation (8)

$$Y_{ij} = \mu_{lj} + \pi_j * T_{ij} + \beta_{1lj} * (X_{ij} - c_j) + \beta_{2lj} * (X_{ij} - c_j)^2 + (\beta_{1rj} - \beta_{1lj}) * T_{ij} * (X_{ij} - c_j) + (\beta_{2rj} - \beta_{2lj}) * T_{ij} * (X_{ij} - c_j)^2 + \upsilon_{ij} \quad (8)$$

## Data

The estimation is conducted combining two sources of data. First, administrative data from the Ministry of Housing of Chile generates the necessary information about applicants for housing units between the years 1998 and 2003 for the three largest regions of Chile[4]. From this data set, it is possible to recover the relevant cutoff for each call made by the Ministry of Housing during this period and the score obtained by each applicant. Second, data collected by the Ministry of Social Planning between November 2006 and April 2007 generates information about outcomes variables. This survey is called "Ficha de Protección Social" and replaced the existing score card used to target social programs (nota al pie: FICHA CAS). Both data sets are matched by the Ministry of Housing for the purpose of this study and a total sample of 222,141 observations from 664 public calls is provided.

For each observation there is an individual score obtained in the application, applicants are ranked by their score and the cutoff is calculated here by setting it equal to the score obtained by the last winner in each public call. The running variable is then constructed by taking the difference between the individual score and the relevant cutoff.

One of the most common threats for the validity of the regression discontinuity approach is the manipulation of the running variable. This is a problem because the basic underlying assumption, continuity of the conditional expectation of counterfactual outcomes in the running variable, is not credible if agents can allocate endogenously in one side of the threshold. A first and simple way to evaluate this issue is by looking at the density of the running variable in each side of the cutoff.

Figure 1A present a histogram using a bandwidth equal to one for those observations fifty points away from the cutoff and shows a clear discontinuity at the threshold. In the context of the assignment of the treatment, there are two reasons that can explain this concentration of density just above the threshold. First, in the housing programs described above applicants are allowed to submit their applications individually or in groups. Cutoff are generated independently in each process, and by definition in the case of the groups a concentration of people is going to be sitting over the threshold. Second, after a close look at the data I found that in many public calls everyone receives the subsidy.

[4] Regions RM, V and VIII.

I address this issue in a more systematic way by implementing the two-step procedure proposed by McCrary (2008). The first step consists of estimating a normalized frequency of the forcing variable using small and equally spaced bins. In the second step, the histogram is smoothed using local linear regression separately on each side of the threshold and frequency is treated as the dependent variable. This is a formal way to test for a discontinuity of the dependent variable in the cutoff. In this setting we can treat each public call as an independent experiment. Therefore, with the intention of ensuring internal validity only public calls that pass this test are used. In other words, public calls with a significant discontinuity on the density of the running variables around the thresholds are dropped from the sample. The new sample is composed by a total of 107 public calls and 61,989 observations.

**Figure 1: Frequency of the Running Variable (+/- 50)**

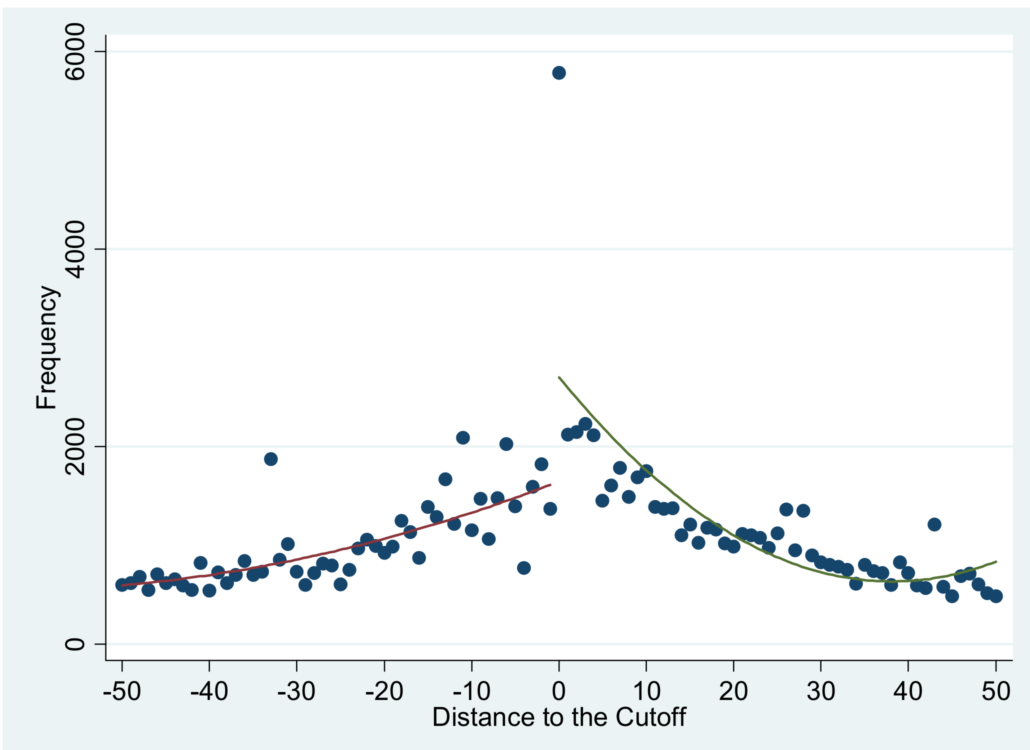
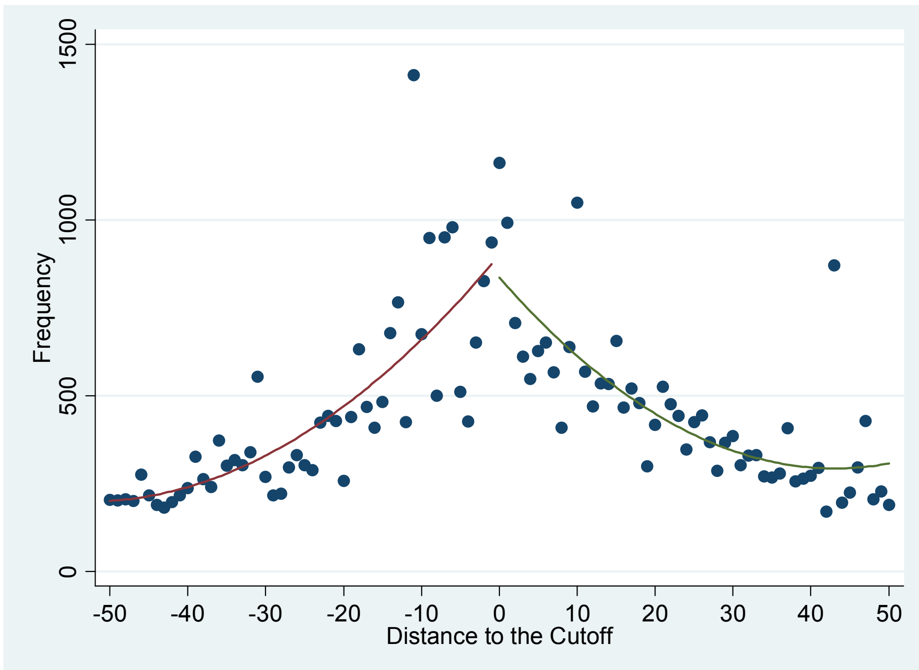


Figure 2 is similar to Figure 1 but uses the new sample. In this case there is no observation of a discontinuity in the density of the running variable around the cutoff, which in turn suggests lack of manipulation for these experiments[5]

**Figure 2: Frequency of the Running Variable for Selected Public Calls (+/-50)**




**Evidence**

I am interested in estimating the impact of the treatment, a publicly provided housing unit, on the level of education achieved by those individual under treatment. An individual will choose to attend one more year of education as long as the marginal benefit of one more year is greater than the corresponding marginal cost. The public provision of a new housing unit can be expected to influence this decision by any of the four elements described above: (1) Income effect: The monetary value of the voucher is about 1.5 to 3 times the annual income of an eligible family, and as suggested by Currie and Yelowitz (2000) this may induce treated families to consume more education. If suddenly the family doesn't have to pay rent, the opportunity cost of getting an extra year of education may go down, (2) Housing effect: Better housing conditions may generate better environment for education (e.g. more space, warmer environment), (3) Neighborhood effect: The

[5] The formal test using the entire sample also rejects the hypothesis of manipulation.

new neighborhood where the housing unit is going to be located may have better, similar of worst conditions than the previous one and this could have a direct influence in educational outcomes. (4) Homeownership effect: the treatment may induce early adoption of homeownership, which according to Glaeser and DiPasquale (1998) may affect education through social capital improvements, or by reducing residential mobility (Aaronson, 2000).

In what follows, I will estimate the overall effect of the treatment on years of education, and although I am unable to identify the relative importance of these four channels, I do evaluate whether each one of them is playing a role induced by the treatment.

## Graphical Analysis

Before showing the estimated coefficients it is useful to provide a visual representation of the relevant variables. This is a simple way of visualizing the functional form of the regression at each side of the cutoff, which in turn gives us a sense of how good approximation a local linear regression and a polynomial can be under different specifications. An additional advantage of the graphical analysis is that we can get a first impression of the size of the jump in the outcome at the threshold. In general, it is recommended to under-smooth the graph using smaller bandwidths in order to see the raw data before imposing any assumptions. I will use $h$=9 and $h$=6 in this section, which are smaller bandwidths compared to the $h$*=18 recommended by the cross validation function in the next section.

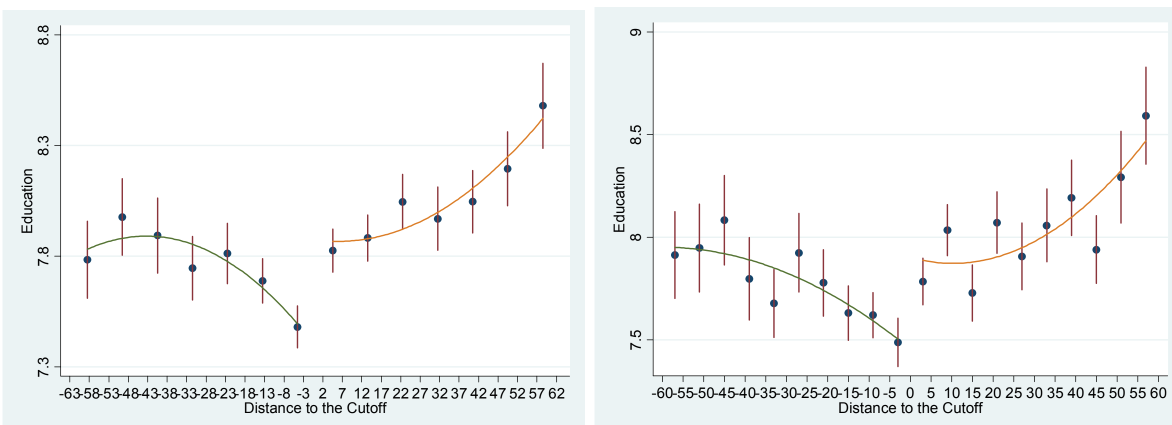
I use years of education as the relevant outcome. The first visualization has this outcome on the vertical axis and the difference between the score of the applicant and the cutoff on the horizontal axis. Thus, the running variable is positive as long as the applicant gets the vouchers, and negative otherwise. I first divide the running variable into several bins leaving aside observations further than 60 points from the cutoff. Using $h$ equal 6 and 9 as the bandwidths I obtain a total of 20 bins and 14 bins, respectively.

The average of the outcome variable is computed for each bin

$$\overline{Y_k} = \frac{1}{N}\sum_{i=1}^{N} Y_i * 1(b_k < X_i < b_{k+1}) \qquad (9)$$

Figure 4 plots this estimate on the vertical axis and the running variable on the horizontal axis pooling all public calls together.[6] On the left side I present the graph using a bandwidth equal to 9, and on the right side one equal to 6. In both cases it is ease to observe that the level of education jumps at the threshold. As a first approximation, the discontinuity of the quadratic fitted line suggests that being offered a housing voucher is associated with an increase in approximately 0.2 years of education, being this our first approximation to $\alpha_{ITT}$, the ¨Intent-to-Treat¨ effect of the housing program.

**Figure 3: Years of Education and Distance to the Cutoff**




Bandwidth=9 Bandwidth=6

Going from the "Intent-to-Treat" to the "Treatment-on-Treated" requires some extra information. In particular I need an estimate for the denominator of equation (4): $E[D_{TOT}|X \geq c] - E[D_{TOT}|X \leq c]$ for observations close to the cutoff. As described before, I am forced to use aggregate take-up information provided by the Ministry of Housing as a proxy for this parameter. [7]

I compute the average take-up rate for the two programs considered here and then I take the weighted average of these two figures using the proportion of individuals that belong to each program in the data set. This figure turns out to be 77%, and a detailed description of the information provided by the Ministry of Housing can be found in appendix A.3. Thus, according

[6] As each public call corresponds to a different experiment, it is not completely correct to pool them together, but the graphical analysis is a useful technique to get a general sense of the shape of the function and the magnitude of the parameter.
[7] Figures are available for download in www.observatoriohabitacional.cl

to our first approximation to the ¨Intent-to-Treat¨ the ¨Treatment-on-Treated¨ becomes

$\alpha_{TOT} = \frac{0.2}{.77} = 0.26$ years of education.

## Local Linear Regression

In this section I formalize the graphical analysis using local linear regressions; I run equation (7) adding fixed effect for each public call *j* in order to recover a weighted average of the parameters that we would have obtained by running the same specification for each public call separately. The key remaining step required to run this local linear regression is the selection of the width of the bins. Section 3 provides a detailed description of the tradeoff between bias and precision, but the basic intuition is as follows: while large bins make the comparison above and below the cutoff less credible, narrowing bins to the extreme generate noise estimates

Here I use a standard ¨leave on out¨ cross-validation procedure for choosing the optimal bandwidth, and since we are interested in the parameter at the threshold I follow Imbens and Lemieux (2007) ´s suggestion of dropping 50% of the observations at the extremes of each side of the cutoff before running the cross-validation procedure. Figure 4 presents the value of the cross validation function for a range of different bandwidths, and the minimum value of the function is obtained using a bandwidth equal to 18 points. This is going to be the preferred bandwidth, but with the intention of evaluating the stability of the coefficients in what follows I will present the estimated coefficients using three different bandwidths: 9, 18 and 27.

**Figure 4: Cross Validation Function**

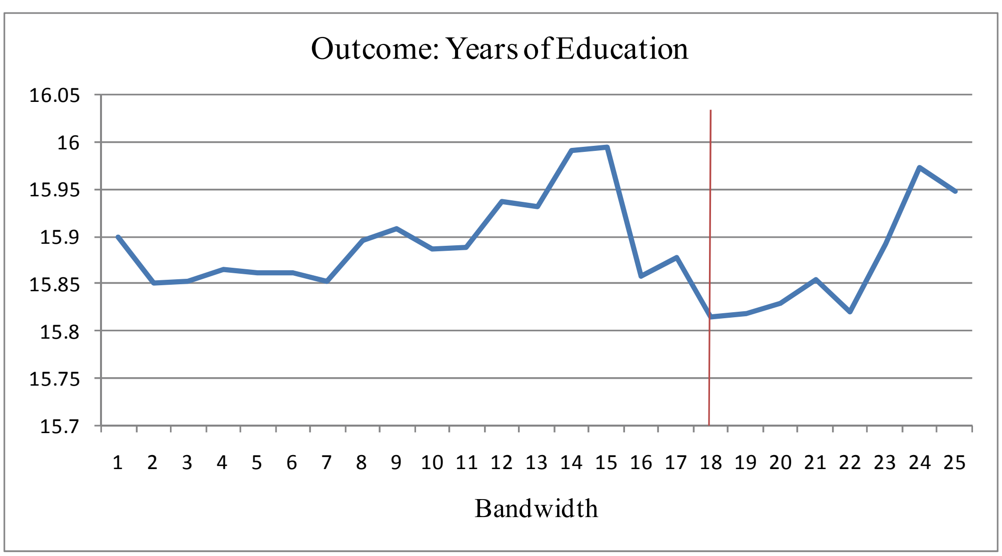


Table 1 confirms what the graphical analysis suggested before. The parameter of interest for the preferred bandwidth has a point estimate equal to 0.211 and it is statistically significantly different from zero at the five percent level. I obtain this estimate of $\alpha_{ITT}$ running a local linear regression similar to equation (7), with years of education as the dependent variable, a bandwidth of 18 as the preferred alternative, and fixed effects for each public call. This suggests that the local ITT of providing a housing unit on education is equivalent to an increase of 0.211 years of education. This result remains unchanged when I increase the bandwidth up to *30*, but the coefficient is not statistically significantly different from zero when the size of the bandwidth is 10.

Panel B of table presents the second coefficient of interest, the effect of the voucher for those who actually used it. Plugging the 77% take up rate in the denominator of equation (4) the 0.221 is amplified by (1/0.77) and the preferred specification suggest that the using a housing voucher of this type increases the level of education in 0.27 years.

One may be worried that this result could be somewhat biased since the significance of the coefficient disappear when we reduce the size of the bandwidth. As a way to check for the robustness of the result for the preferred bandwidth it is useful to compare these non parametric estimates with polynomial regression analysis.

**Polynomial Regressions**

I repeat the previous analysis using a polynomial specification as a way to evaluate the consistency of the estimates obtained using a non parametrical approach. Instead of selecting the optimal bandwidth, the relevant point here is the consistency of the parameters using different orders of the polynomial function. I report the coefficients for second, third and fourth order polynomial regressions. Table 2 presents coefficients that are stable across the different specifications, and consistent with the evidence found in Table 1. For $\alpha_{ITT}$ the point estimate is consistently suggesting that receiving the housing voucher increases education by about 0.15 years, and these coefficients are statistically significantly different from zero. The corresponding $\alpha_{TOT}$ point estimate suggest that actually using the voucher and getting a new housing unit increases education in 0.2 years. The economic significance of a point estimate of 0.2 is substantial considering that so far we have not explored the possibility of a differential effect for a particular subgroup of the population.

I explore this alternative running gender and age subgroup analysis. For the gender subgroup analysis I create a dummy variable called $M_{ij}$ that takes value equal one if the individual is male, and zero otherwise. For the age decomposition I create a dummy variable called $A_{ij}$ that takes value equal to one if the individual's age is more than 25 years old, assuming that most of the effect is not going to be concentrated among this subgroup of the population.

TABLE 1: LOCAL LINEAR REGRESSION

| Dependent Variable: Years of Educaction | Bandwidth +/- 18 | Bandwidth +/- 10 | Bandwidth +/- 30 |
|---|---|---|---|
| *Panel A* : ITT | (1) | (2) | (3) |
| *Average Years of Education* | 8.37 | 8.33 | 8.30 |
| *(s.d)* | (3.96) | (3.98) | (3.96) |
| *Regressors* | | | |
| *Intend-to-Treat ($T_{ij}$)* | 0.211** | 0.117 | 0.206*** |
| | (0.097) | (0.125) | (0.079) |
| *Running Variable ($X_{ij}$-$c_j$)* | -0.002 | -0.001 | -0.005 |
| | (0.007) | (0.016) | (0.004) |
| *Interaction ($X_{ij}$-$c_j$)*$T_{ij}$* | -0.005 | 0.017 | 0.001 |
| | (0.010) | (0.022) | (0.006) |
| *Panel B:* TOT | (1) | (2) | (3) |
| *Trake up rate* | 0.77 | 0.77 | 0.77 |
| *Treatment-on-Treated ($\alpha_{TOT}$)* | 0.274 | 0.152 | 0.268 |
| Observations | 23,679 | 13,903 | 32,224 |

Notes: Regression with fixed effects for each public call where a cutoff is generated. Robust standard errors clustered at the individual level are shown in parentheses. * Significant at the 10-percent level ** Significant at the 5-percent level *** Significant at the 1-percent level. TOT in Panel B is estimated by multiplying the Intend-to-Treat coefficient by the inverse of the take up rate.

Table 3 presents the subgroup analysis using the preferred non parametrical (bandwidth=18) and parametric third order polynomial specification. For the gender subgroup analysis I find mix evidence. On the one hand I find that the negative differential effect of the treatment on men is not statistically different from zero using local linear regression. On the other hand the third order polynomial regression suggests a higher magnitude for the negative differential effect for men, although the coefficient is only significant at the 10 percent level.

As we may have expected, evidence on age is more robust. Under both specifications the point estimate for the ¨Intent-to-Treat¨ effect goes up to about 0.3 years of education for the population under 25 years old. This parameter is statistically significantly different from zero at the five percent level in both cases. The point estimate of the differential effect for the adult population is -0.14 for the local linear regression and -0.21 for the polynomial specification, although only later coefficient statistically significantly different from zero.

| TABLE 2: POLYNOMIAL REGRESSION | | | |
|---|---|---|---|
| Dependent Variable: Years of Educaction | Specification | | |
| | Second Order | Third Order | Fourth Order |
| *Panel A* : ITT | (1) | (2) | (3) |
| *Average Educacion* | 7.915 | 7.915 | 7.915 |
| *(s.d)* | (3.856) | (3.856) | (3.856) |
| *Regressors* | | | |
| *Intend-to-Treat ($T_{ij}$)* | 0.159*** | 0.145** | 0.144* |
| | (0.0588) | (0.0704) | (0.0821) |
| *Running Variable ($X_{ij}$-$c_j$)* | 0.000184 | 0.00105 | -0.00154 |
| | (0.00264) | (0.00511) | (0.00877) |
| *Interaction ($X_{ij}$-$c_j$)*$T_{ij}$* | 0.000481 | 6.73e-05 | 0.00389 |
| | (0.00281) | (0.00538) | (0.00914) |
| *2nd Order Running Variable ($X_{ij}$-$c_j$)^2* | 3.72e-05 | 5.72e-05 | -3.98e-05 |
| | (2.42e-05) | (0.000103) | (0.000292) |
| *2nd Order Interaction ($X_{ij}$-$c_j$)^2*$T_{ij}$* | -4.13e-05* | -6.51e-05 | 1.44e-05 |
| | (2.43e-05) | (0.000103) | (0.000293) |
| *3rd Order Running Variable ($X_{ij}$-$c_j$)^3* | | 1.15e-07 | -1.08e-06 |
| | | (5.76e-07) | (3.48e-06) |
| *3rd Order Interaction ($X_{ij}$-$c_j$)^3*$T_{ij}$* | | -1.09e-07 | 1.16e-06 |
| | | (5.76e-07) | (3.48e-06) |
| *4th Order Running Variable ($X_{ij}$-$c_j$)^4* | | | -4.59e-09 |
| | | | (1.34e-08) |
| *4th Order Interaction ($X_{ij}$-$c_j$)^4*$T_{ij}$* | | | 4.52e-09 |
| | | | (1.34e-08) |
| *Panel B:* TOT | (1) | (2) | (3) |
| *Take up rate* | 0.77 | 0.77 | 0.77 |
| *Treatment-on-Treated ($\alpha_{TOT}$)* | 0.206 | 0.188 | 0.187 |
| Observations | 61,989 | 61,989 | 61,989 |

Notes: Regression with fixed effects for each public call where a cutoff is generated. Robust standard errors clustered at the individual level are shown in parentheses. * Significant at the 10-percent level ** Significant at the 5-percent level *** Significant at the 1-percent level. TOT in Panel B is estimated by multiplying the Intend-to-Treat coefficient by the inverse of the take up rate.

In Panel B of Table 3 we can observe that the impact of actually using the housing voucher for the young population, the ¨Treatment-on-Treated¨, is 0.39 years for the nonparametric and 0.353 years for the parametric specification. Both co-efficient are significant at the 5 percent level.

TABLE 3: SUBGROUP ANALYSIS

| Dependent Variable: Years of Educaction | Specification | |
|---|---|---|
| | LLR Bandwidth=18 | Polynomial Third Order |
| *Panel A* : ITT | (1) | (2) |
| *Average Educacion* | 8.367 | 7.915 |
| *(s.d)* | (3.958) | (3.856) |
| *Gender Subgroup Analysis* | | |
| *Intend-to-Treat ($T_{ij}$)* | 0.230** | 0.198*** |
| | (0.104) | (0.0762) |
| *Differential effect for men ($T_{ij}$*$M_{ij}$)* | -0.031 | -0.111* |
| | (0.155) | (0.0629) |
| *Age Subgroup Analysis* | | |
| *Intend-to-Treat ($T_{ij}$)* | 0.302** | 0.272*** |
| | (0.145) | (0.0855) |
| *Differential effect for adult population ($T_{ij}$*$A_{ij}$)* | -0.138 | -0.206*** |
| | (0.210) | (0.0676) |
| *Panel B:* TOT | (1) | (2) |
| *Take up rate* | 0.77 | 0.77 |
| *Women-Treatment-on-Treated ($\alpha_{TOT}$)* | 0.299 | 0.257 |
| *Young-Treatment-on-Treated ($\alpha_{TOT}$)* | 0.391 | 0.353 |
| Observations | 23,679 | 61,989 |

Notes: Regression with fixed effects for each public call where a cutoff is generated. adult population is defined by those with more than 25 years old. Robust standard errors clustered at the individual level are shown in parentheses. * Significant at the 10-percent level ** Significant at the 5-percent level *** Significant at the 1-percent level. TOT in Panel B is estimated by multiplying the Intend-to-Treat coefficient by the inverse of the take up rate.

## Housing Effect

What is driving the previous finding? Is it the income effect? Is it better housing conditions? Is it a better neighborhood? Or it is the homeownership? Here I see whether these channels are activated by the treatment. For the case of the income effect we know, by definition, that is part of the treatment. We also know that households in the control group do not get this element of the treatment. This is not necessarily the case for the other three components. It is possible, for example, that the control group obtains similar housing conditions by their own means. The same could be true for neighborhood and homeownership conditions.

In terms of the housing conditions I look at four variables: (1) bedroom density, which is computed dividing the number of people living in the household by the number of bedrooms in the house, (2) sewer system (3) in-house drinkable water, and (4) dirt floor in the housing unit. As usual the information comes from a survey taken between three and six years after the voucher was granted.

Figure 5 presents the graphical analysis for these four dimensions. For the graphical I start from the bandwidth recommended by the cross validation function and then iterate reducing its size until the noise starts interfering with the interpretation, which happens between h=3 and h=6.

From the graphs it is apparent that the treatment is positively affecting three of the four dimensions analyzed here, being "Dirt Floor" the exception. Using the discontinuity in the graph as a way to get a first approximation to the ¨Intent-to-Treat¨ effect, for the case of Bedroom Density we observe a reduction of about 0.1 people per room, going from about 2.05 to 1.93. The probability of having a Sewer System is also affected and goes up by about 10 percentage points, from 0.58 to about 0.68, and the probability of having in-house drinkable water goes up by about 9 percentage points, from 0.73 to 0.82.

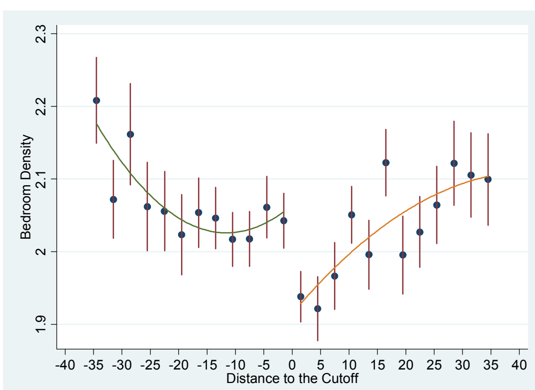
The baseline probability of living in a house with Dirt Floor is about 0.6%, which suggests this type of treatment is not targeted to those living in extreme conditions. Nonetheless, this is the only variable considered here that it doesn't seems to be affected by the treatment. A plausible hypothesis for this feature, although not possible to test here, is that those who used to live in extreme conditions are less likely to use the voucher. This would be compatible with evidence from the field suggesting that for the poorest households getting a new housing unit is associated with paying bills for the first time.

Table 4 provides the formal analysis of the graphical analysis, and most of the findings are confirmed, particularly for the more flexible specifications, or when using smaller bandwidths. In terms of the ¨Treatment-on-Treated¨, and using the Local Linear Regression specification with a bandwidth equal to 18  I find that using the housing voucher decreases the number of people per bedroom in 0.19 for which is equivalent to a 10% reduction in the parameter; the probability of having a sewer system increases in 11 percentage points, equivalent to an increase of about 14%; the probability of having in-house drinkable water increases  by 7 percentage points,  or a 10% increase  in this variable. All of these coefficients are statistically significantly different from zero at the one percent level.

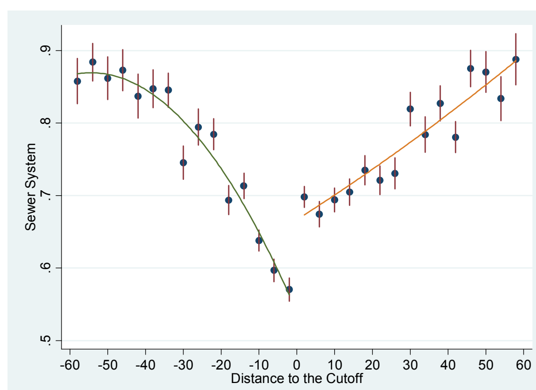
The last variable of interest for the housing effect, the probability of living in a housing unit with Dirt Floor, has a point estimate is positive and statistically significantly different from zero for the treatment, but the economic significance of the coefficient is minor (0.0037), which is consistent with Figure 5c.

**Figure 5: Housing Effect**

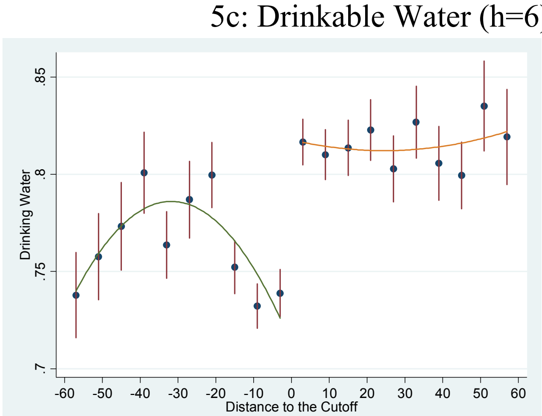
5a: Bedroom Density (h=3)



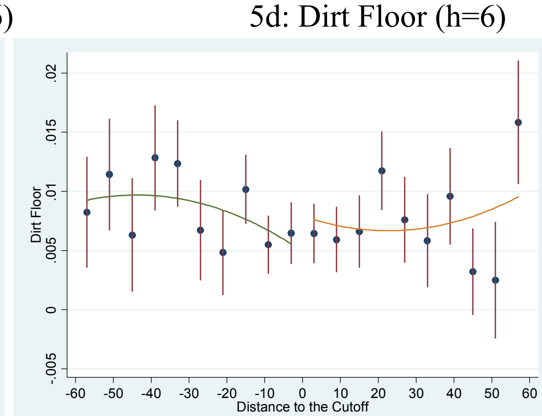

5b: Sewer System (h=4)




5c: Drinkable Water (h=6)




5d: Dirt Floor (h=6)



Comparing column (1) and column (6) from Table 4 is an easy way to evaluate the stability of the nonparametric coefficients with respect the parametric specification. Column (1) presents the relevant parameter using a bandwidth equal to 18 for the Local Linear Regression and column (6) is the most flexible of the parametric specifications. The general finding is that the coefficient has a similar magnitude and sign.

The conclusion from this analysis is that the treatment under consideration has a positive and significant effect on the characteristics of housing units used by the winners of the voucher. Thus, it is possible to speculate that the housing effect is contributing to the extra 0.39 years of education generated by the treatment for the population under 25.

TABLE 4: HOUSING EFFECT

| | LOCAL LINEAR REGRESSION | | | POLYNOMIAL REGRESSION | | |
|---|---|---|---|---|---|---|
| | Bandwidth | | | Specification | | |
| | +/- 18 | +/- 10 | +/- 30 | 2nd Order | 3rd Order | 4th Order |
| *Panel A* : ITT | (1) | (2) | (3) | (4) | (5) | (6) |
| *Impact on Bedroom Density* | -0.150*** | -0.143*** | -0.113*** | -0.0266* | -0.0690*** | -0.0994*** |
| average = 2.087 | (0.0217) | (0.0276) | (0.0182) | (0.0140) | (0.0166) | (0.0194) |
| *Impact on Sewer System* | 0.0852*** | 0.132*** | 0.0609*** | 0.0263*** | 0.0406*** | 0.0616*** |
| average = 0.768 | (0.0102) | (0.0132) | (0.00827) | (0.00596) | (0.00723) | (0.00848) |
| *Impact on Drinkable Water* | 0.0506*** | 0.0578*** | 0.0681*** | 0.0458*** | 0.0578*** | 0.0635*** |
| average = 0.784 | (0.00998) | (0.0128) | (0.00817) | (0.00610) | (0.00737) | (0.00860) |
| *Impact on Dirt Floor* | 0.00387** | 0.00352* | 0.000717 | 0.00161 | 0.00285** | 0.00317** |
| average = 0.0067 | (0.00192) | (0.00205) | (0.00148) | (0.00107) | (0.00128) | (0.00146) |
| *Panel B:* TOT | (1) | (2) | (3) | (4) | (5) | (6) |
| *Take up rate* | 0.770 | 0.770 | 0.770 | 0.770 | 0.770 | 0.770 |
| *Impact on Bedroom Density* | -0.195 | -0.186 | -0.147 | -0.035 | -0.090 | -0.129 |
| *Impact on Sewer System* | 0.111 | 0.171 | 0.079 | 0.034 | 0.053 | 0.080 |
| *Impact on Drinkable Water* | 0.066 | 0.075 | 0.088 | 0.059 | 0.075 | 0.082 |
| *Impact on Dirt Floor* | 0.005 | 0.005 | 0.001 | 0.002 | 0.004 | 0.004 |
| Observations | 23,679 | 13,903 | 32,224 | 61,989 | 61,989 | 61,989 |

Notes: Regression with fixed effects for each public call where a cutoff is generated. adult population is defined by those with more than 25 years old. Robust standard errors clustered at the individual level are shown in parentheses. * Significant at the 10-percent level ** Significant at the 5-percent level *** Significant at the 1-percent level. TOT in Panel B is estimated by multiplying the Intend-to-Treat coefficient by the inverse of the take up rate.

**Neighborhood Effect**

The survey taken between three and six years after the treatment has information about the municipality where the housing unit is located, which in turn allow us to compute some environmental variables for their neighborhood. Information about the characteristic of the municipalities is drawn from CASEN 2006, a nationally representative survey commissioned by the Ministry of Social Planning.

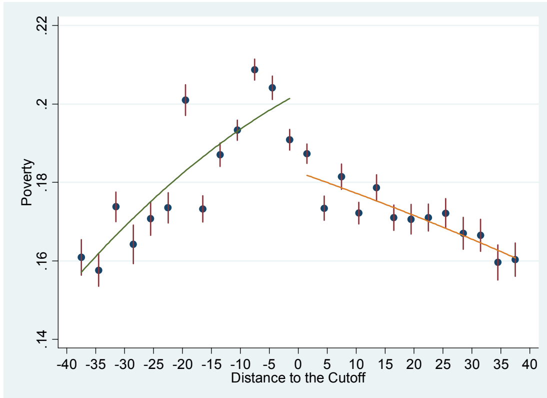
With this information I can test whether the mobility component, the fact that new housing units are normally located in the perimeter of the city, is actually moving the beneficiaries into a worst neighborhood that could be offsetting the effect of having better housing conditions. The unit of analysis is the municipality, and I am going to evaluate the impact of the treatment on three neighborhood variables: (1) Poverty levels computed using the headcount ratio, (2) Delinquency levels computed using crime reports per 100,000 habitants, and (3) Rate of unemployment. In addition, and because unemployment in the municipality may be a bad proxy for the labor constrains faced by the individual after changing places, I also present data on individual levels of employment. This is done using a dummy variable indicating whether the person is working as the outcome variable.[8]

Figure 6 report the graphical analysis for these environmental outcomes. The general conclusion is that there are not significant differences between the treatment and the control group, and if anything, it looks that poverty levels are slightly lower in the new neighborhoods.

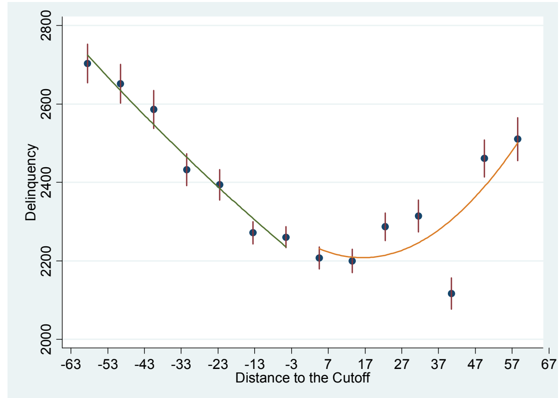
[8] Note that the level of employment may also be affected by the impact of the income affect over labor supply.

**Figure 6: Neighborhood Effect**

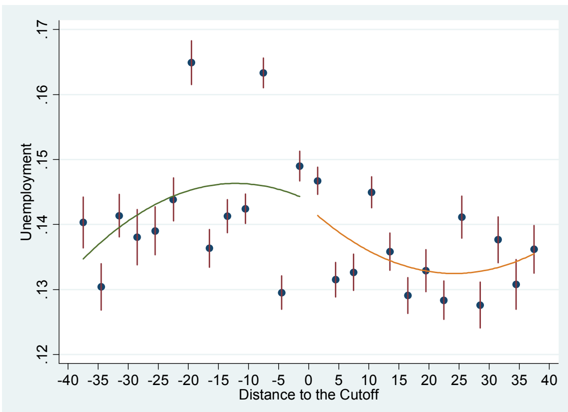
6a: Level of Poverty (h=3)




6b: Delinquency (h=9)




6c: Unemployment (h=3)




6d: Individual Employed (h=3)

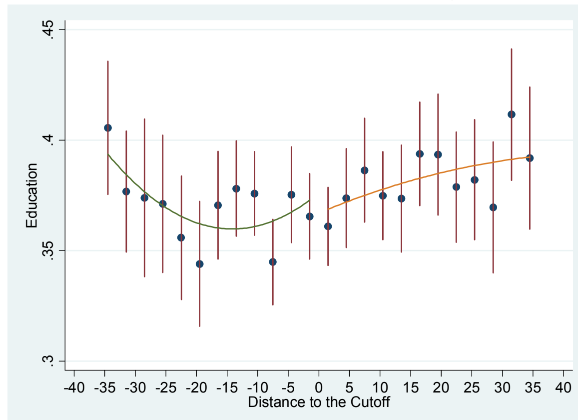



Table 5 present the regression analysis for the neighborhood effect, and consistent with the graphical analysis we find that if anything, those who receive the voucher live in neighborhoods with lower levels of poverty, although the difference is very small, and from the graphical analysis is not clear whether the estimated discontinuity is part of the functional form of the curve. For Delinquency and Unemployment the point estimates are economically negligible, and for the case of the individual level of employment there is no evidence of discontinuity at all.

| TABLE 5: NEIGHBORHOOD EFFECT | | | | | | |
|---|---|---|---|---|---|---|
| | LOCAL LINEAR REGRESSION | | | POLYNOMIAL REGRESSION | | |
| | Bandwidth | | | Specification | | |
| | +/- 18 | +/- 10 | +/- 30 | 2nd Order | 3rd Order | 4th Order |
| *Panel A* : ITT | (1) | (2) | (3) | (4) | (5) | (6) |
| *Impact on Poverty* [1] | -0.0120*** | -0.00356** | -0.00798*** | -0.00192*** | -0.00403*** | -0.00817*** |
| average = 0.171 | (0.00116) | (0.00156) | (0.000974) | (0.000742) | (0.000881) | (0.00103) |
| *Impact on Delinquency* [2] | -77.85*** | 91.97*** | -124.9*** | -147.2*** | -136.6*** | -98.68*** |
| average = 2,380 | (21.48) | (25.97) | (17.05) | (14.28) | (16.33) | (18.56) |
| *Impact on Unemployment* | -0.00320*** | -0.000736 | 0.00317*** | 0.00138** | 0.00115 | 0.000167 |
| average = 0.138 | (0.000992) | (0.00132) | (0.000822) | (0.000577) | (0.000713) | (0.000855) |
| *Impact on Indivual Employment* | -0.00486 | -0.0150 | -0.00505 | 0.00255 | 0.000587 | -0.0001 |
| average = 0.377 | (0.0117) | (0.0151) | (0.00954) | (0.00712) | (0.00852) | (0.00993) |
| *Panel B:* TOT | (1) | (2) | (3) | (4) | (5) | (6) |
| *Take up rate* | 0.770 | 0.770 | 0.770 | 0.770 | 0.770 | 0.770 |
| *Impact on Poverty* | -0.016 | -0.005 | -0.010 | -0.002 | -0.004 | -0.011 |
| *Impact on Delinquency* | -101.1 | 119.4 | -162.2 | -191.2 | -177.4 | -128.2 |
| *Impact on Unemployment* | -0.004 | -0.001 | 0.004 | 0.002 | 0.001 | 0.000 |
| *Impact on Indivual Employment* | -0.006 | -0.015 | -0.007 | 0.003 | 0.001 | 0.000 |
| Observations | 23,679 | 13,903 | 32,224 | 61,989 | 61,989 | 61,989 |

Notes: Regression with fixed effects for each public call where a cutoff is generated. adult population is defined by those with more than 25 years old. Robust standard errors clustered at the individual level are shown in parentheses. * Significant at the 10-percent level ** Significant at the 5-percent level *** Significant at the 1-percent level.  TOT in Panel B is estimated by multiplying the Intend-to-Treat coefficient by the inverse of the take up rate. (1) Poverty mesured using Headcount Ratio (2) Rate of crime reports per 100,000 habitants

## Homeownership Effect

It is natural to expect that the treatment would have an effect on homeownership, but the interpretation of this effect is not completely clear in this context.  Figure 7 plots the rate of homeownership and the running variables using two different bandwidths. On the left side a bandwidth equal to 3, and on the right side a bandwidth three times larger. In both cases it is clear that the probability of being homeowner jumps at the threshold, and from the discontinuity of the flexible fitted line we get a sense that the magnitude is slightly higher than 10 percentage points.

**Figure 7: Homeownership Effect**

7a: Bandwidth=3 7b: Bandwidth=9

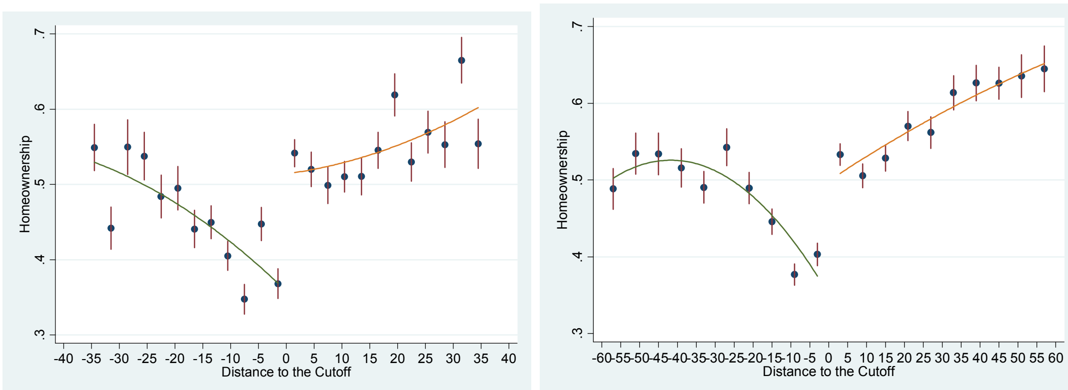



Table 6 confirms what the graphical analysis suggested. The average rate of homeownership is 0.52 for the entire sample, and the discontinuity at the threshold has a point estimate equal to 11 percentage points for the preferred bandwidth. This estimate is statistically significantly different from zero at the 1 percent level and the result is relatively stable for the different specifications.

TABLE 6: HOMEOWNERSHIP EFFECT

| | LOCAL LINEAR REGRESSION | | | POLYNOMIAL REGRESSION | | |
|---|---|---|---|---|---|---|
| | Bandwidth | | | Specification | | |
| | +/- 18 | +/- 10 | +/- 30 | 2nd Order | 3rd Order | 4th Order |
| *Panel A* : ITT | (1) | (2) | (3) | (4) | (5) | (6) |
| *Impact on Homeownerhip* | 0.113*** | 0.139*** | 0.111*** | 0.0858*** | 0.0951*** | 0.101*** |
| average = 0.527 | (0.011) | (0.015) | (0.009) | (0.00707) | (0.00839) | (0.00971) |
| *Panel B:* TOT | (1) | (2) | (3) | (4) | (5) | (6) |
| *Take up rate* | 0.770 | 0.770 | 0.770 | 0.770 | 0.770 | 0.770 |
| *Impact on Homeownerhip* | -0.016 | -0.005 | -0.010 | -0.002 | -0.004 | -0.011 |
| Observations | 23,679 | 13,903 | 32,224 | 61,989 | 61,989 | 61,989 |

Notes: Regression with fixed effects for each public call where a cutoff is generated. adult population is defined by those with more than 25 years old. Robust standard errors clustered at the individual level are shown in parentheses. * Significant at the 10-percent level ** Significant at the 5-percent level *** Significant at the 1-percent level. TOT in Panel B is estimated by multiplying the Intend-to-Treat coefficient by the inverse of the take up rate. (1) Poverty mesured using Headcount Ratio (2) Rate of crime reports per 100,000 habitants

The evidence presented above suggests that the public provision of a housing unit targeted to the low income population of Chile has a positive and significant effect on the level of education acquired by the young population of the households under treatment.

Among the mechanisms that could explain this phenomenon, it is clear that neighborhood effects are not activated by the treatment, and it seems that the Homeownership channel is playing a secondary role. If anything, we would expect that Homeownership is going to have an impact in the longer run through its positive externality on social capital, or by reducing residential mobility. On the contrary, is not hard to think that an important income transfer and better housing conditions could interact with education.

## Conclusion

In this study I estimate the impact of publicly provided new housing units on education by taking advantage of a targeting system for housing vouchers implemented in Chile in the early eighties. The impact of the treatment is estimated by comparing the outcome for applicants with scores just above and below the threshold generated in each assignment process.

I estimate the impact of publicly provided homeownership on the level of education obtained by the population under 25 years old, and after three to six years I find that the treatment increased by 0.39 years the level of education. There is also evidence, although less robust to different specification, that the benefit is concentrated among women. Among the mechanisms that could explain this phenomenon, it is clear that neighborhood effects are not activated by the treatment, and it seems that the Homeownership channel is playing a secondary role. If anything, we would expect that Homeownership is going to have an impact in the longer run through its positive externality on social capital, or by reducing residential mobility. On the contrary, is not hard to think that an important income transfer and better housing conditions could interact with education.

These results are consistent with the positive income effect on education found by Currie and Yelowitz (2000) and also provide valuable information about a relevant channel that may affect educational outcomes in developing countries. While these results are also consistent with current evidence in developed countries suggesting that promoting homeownership has a positive impact on education by reducing residential mobility (Aaronson, 2000) , or by generating social capital (Glaeser and DiPasquale, 1998), the magnitude of the effect suggests that in developing countries education can be directly constrained by precarious housing conditions and income needs.

**References**


Aaronson, D. (2000). A Note on the Benefits of Homeownership. Journal of Urban Economics, 47(3), 356-369.

Buckley, R.M., & Kalarickal, J. (2006). Thirty years of World Bank shelter lending: what have we learned? : World Bank Publications.

Currie, J., & Yelowitz, A. (2000). Are public housing projects good for kids? Journal of Public Economics, 75(1), 99-124.

Dietz, R.D., & Haurin, D.R. (2003). The social and private micro-level consequences of homeownership. Journal of Urban Economics, 54(3), 401-450.

DiPasquale, D., Glaeser, E.L., & National Bureau of Economic, R. (1998). Incentives and social capital: are homeowners better citizens? : National Bureau of Economic Research Cambridge, Mass., USA.

Fisher, L.M., & Jaffe, A.J. (2003). Determinants of international home ownership rates. Housing finance international, 18(1), 34-37.

Gilbert, A. (2004). Helping the poor through housing subsidies: lessons from Chile, Colombia and South Africa. Habitat International, 28(1), 13-40.

Green, R.K., & White, M.J. (1997). Measuring the benefits of homeowning: Effects on children. Journal of Urban Economics, 41(3), 441-461.

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

Haurin, D.R., Parcel, T.L., & Haurin, R.J. (2002). Does homeownership affect child outcomes? Real Estate Economics, 30(4), 635-666.

Holland, P.W. (1986). Statistics and causal inference. Journal of the American Statistical Association, 945-960.

Imbens, G.W., & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. Journal of Econometrics, 142(2), 615-635.

Jacob, B.A. (2005). Accountability, incentives and behavior: The impact of high-stakes testing in the Chicago Public Schools. Journal of Public Economics, 89(5-6), 761-796.

Katz, L., Kling, J., & Liebman, J.B. (2007). Experimental analysis of neighborhood effects. Econometrica, 75(1), 83–119.

Krugman, P. (2008, June 23). Home Not-So-Sweet Home New York Times.

Lee, D. & Lemieux, T. (2009). Regression Discontinuity Design in Economics. NBE Working Paper 14723

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. Journal of Econometrics, 142(2), 698-714.

Oreopoulos, P. (2003). The Long-Run Consequences of Living in a Poor Neighborhood. Quarterly Journal of Economics, 118(4), 1533-1575.

Pardo, C.A. (1999). Housing Finance in Chile: Primary and Secondary Mortgage Financing. Housing finance international, 14, 15-25.

Retsinas, N. & Belsky, E. (2002). Low-income homeownership: examining the unexamined goal. Joint Center for Housing Studies

Rosen, H. (1985). Housing subsidies. Handbook of public economics, 1, 375-420.

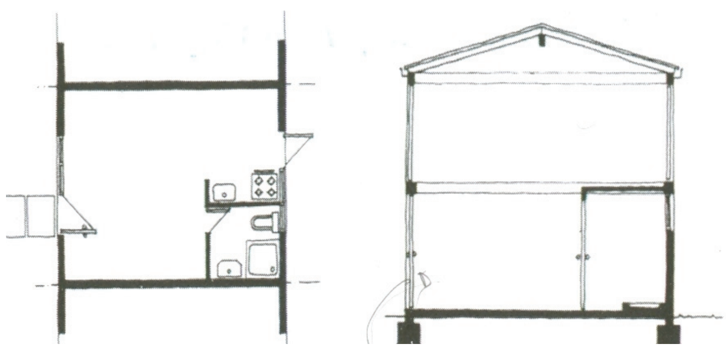
Rubin, D.B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. Journal of Educational Psychology, 66(5), 688-701.

**Appendixes**

## *A.1 Treatment 1: Vivienda Progresiva*

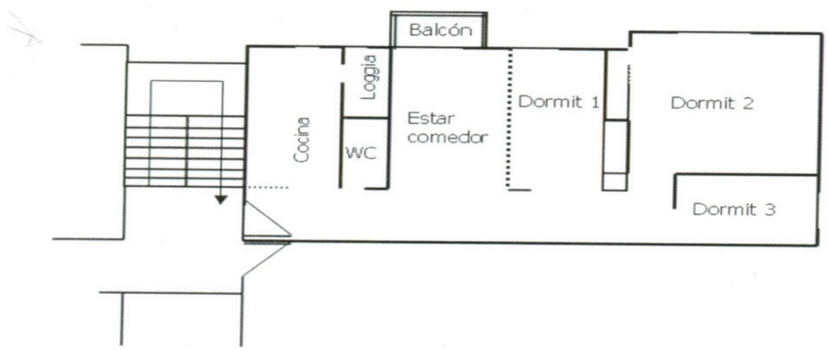
- Site of 100 sq meters with access to water, electricity and sewer system
- Expandable construction of 23 sq meters that includes kitchen, bathroom and one open room
- Value of the product: approximate 6000 USD
- Value of the subsidy: approximate 5700 USD
- Required savings: approximate 300 USD
- The size of the subsidy (6700 USD) is between 2 and 3 times the annual income of an eligible family

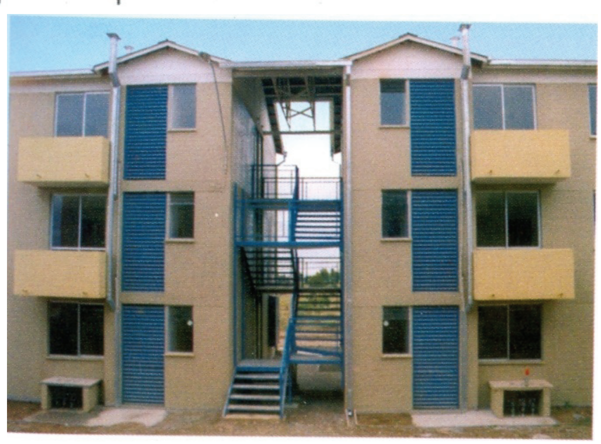
Optional extensions
  - Subsidy between 775$ US and 1500 USD to expand the original construction.
  - Soft credit for additional expansions up to 2000 USD



### *A.2 Treatment 2: Vivienda Básica & Nueva Básica*

- Site of 100 sq meters with access to water, electricity and sewer system
- Apartments or coupled houses with a construction between 38 and 42 sq meters
- Value of the product: approximate 9900 USD
- Value of the Subsidy: approximate 5940 USD
- Required savings: approximate 430 USD
- Credit up to 3440 USD for a period of 8, 10 or 15 years
- The size of the subsidy (5940 USD) is between 1.5 and 3 times the annual income



## ***A.3*** *Number of Housing Subsidies per Year and Program*

**NUMBER OF HOUSING SUBSIDIES GRANTED PER YEAR AND PROGRAM**

| PROGRAM | | YEAR | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | **1990** | **1991** | **1992** | **1993** | **1994** | **1995** | **1996** | **1997** | **1998** | **1999** | **2000** | **2001** | **2002** | **2003** | **2004** | **2005** |
| Viviendas SERVIU contratadas | Vivienda Básica | 18,895 | 25,322 | 25,862 | 26,964 | 25,720 | 24,543 | 23,051 | 19,609 | 24,097 | 21,006 | 21,499 | 21,734 | 6,510 | | | |
| | Vivienda Progresiva | 4,063 | 8,094 | 6,409 | 4,710 | 3,671 | 3,622 | 3,633 | 3,620 | 3,324 | 3,034 | | | | | | |
| Subsidios Otorgados | Nueva Básica | | | | | | 4,972 | 7,132 | 7,298 | 6,115 | 6,574 | 7,334 | 9,826 | 16,216 | 22,049 | 2,068 | 0 |
| | Progresiva I Etapa | | 770 | 4,665 | 7,002 | 8,737 | 6,261 | 10,324 | 8,683 | 9,518 | 9,733 | 9,195 | 9,649 | 8,897 | 9,984 | 4,418 | 1,290 |
| | Progresiva II Etapa | | 250 | 3,395 | 2,010 | 2,428 | 2,238 | 3,103 | 1,567 | 2,360 | 1,885 | 2,142 | 2,062 | 1,816 | 1,155 | 611 | 37 |
| | **TOTAL** | **22,958** | **34,436** | **40,331** | **40,686** | **40,556** | **41,636** | **47,243** | **40,777** | **45,414** | **42,232** | **40,170** | **43,271** | **33,439** | **33,188** | **7,097** | **1,327** |

**NUMBER OF HOUSING SUBSIDIES USED PER YEAR AND PROGRAM**

| PROGRAM | | YEAR | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | **1990** | **1991** | **1992** | **1993** | **1994** | **1995** | **1996** | **1997** | **1998** | **1999** | **2000** | **2001** | **2002** | **2003** | **2004** | **2005** |
| Viviendas SERVIU Terminadas | Vivienda Básica | 15,899 | 24,975 | 22,314 | 22,331 | 24,908 | 27,550 | 25,962 | 19,752 | 14,449 | 21,056 | 20,280 | 20,069 | 21,929 | 8,628 | 5,204 | 1,033 |
| | Vivienda Progresiva | | 6,167 | 8,750 | 5,687 | 4,665 | 3,503 | 4,207 | 2,829 | 3,669 | 3,459 | 2,394 | 232 | 588 | 60 | | |
| Subsidios Pagados | Nueva Básica | | | | | | 753 | 2,555 | 2,198 | 2,827 | 3,398 | 4,701 | 5,698 | 8,204 | 10,890 | 14,002 | 5,546 |
| | Progresiva I Etapa | | 313 | 1,116 | 2,519 | 5,189 | 7,039 | 7,371 | 5,281 | 4,727 | 6,677 | 7,033 | 7,822 | 8,547 | 8,204 | 9,161 | 6,860 |
| | Progresiva II Etapa | | 195 | 111 | 1,581 | 1,450 | 2,226 | 2,582 | 1,703 | 1,217 | 1,242 | 1,319 | 1,714 | 1,774 | 1,254 | 1,038 | 793 |
| **TOTAL** | | **15,899** | **31,650** | **32,291** | **32,118** | **36,212** | **41,071** | **42,677** | **31,763** | **26,889** | **35,832** | **35,727** | **35,535** | **41,042** | **29,036** | **29,405** | **14,232** |

Reference: www.observatoriohabitacional.cl