

Bolivian Social Investment Fund
Analysis of Baseline Data for Impact Evaluation

Menno Pradhan^{*}, Laura Rawlings^{**} and Geert Ridder^{*}

February 1997

(*) Department of Economics, Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands

(**) Policy Research Department, The World Bank, 1818 H street N.W. Washington DC 20433, USA.

The views expressed in this paper are those of the authors. They do not necessarily represent the views of the World Bank.

The analysis presented in this paper was made possible by a grant from the Dutch Trust Fund at the World Bank. The collection of the baseline data was financed by the Social Investment Fund Project (1991-1994). This study is part of the World Bank financed research project "Evaluation of Social Sector Investments" which has provided support to the Bolivian Social Investment Fund for the design, application and analysis of the impact evaluation. The authors wish to thank John Newman and Ramiro Coa for their support.

Abstract:

On the basis of the baseline data collected for the evaluation of the Bolivian Social Investment Fund (SIF) this paper assesses (1) the benefit incidence of the SIF and (2) the quality of the evaluation design. We find that the benefits in education are most equally distributed over the population, the investments in health and sanitation favor those relatively well off. For the education component of the SIF, control groups of schools which will not receive benefits have been included in the survey. In one region these schools were selected on the basis of matched comparison on the basis of observed characteristics, in the other region by means of randomization. We compare control and treatment groups and conclude there is a systematic bias in favor of treatment schools in the first region. We propose to use instrumental variables to control for the non-random selection. With the pre-intervention data we can test whether an instrument is valid. We find that among several candidates the number of NGOs (non governmental organizations) in the community is a valid instrument. Next, we investigate the possible loss of efficiency in the estimate of the impact due to the non experimental control group design.

1. Introduction

The Bolivian Social Investment Fund (SIF) is a financial institution designed to promote sustainable investments in the social sectors, notably in the areas of health, education, and water and sanitation. The SIF co-finances initiatives providing infrastructure, training and equipment by making funds available to requesting agencies who then sub-contract the implementation of the project. The SIF is characterized by rapid disbursement, institutional efficiency and a demand-driven approach, and it is among the first of its kind in the world. The SIF has been a key player in funding social sector projects in Bolivia. Since its inception in 1991, the SIF has provided US\$117 million worth of financing to almost 2,000 projects¹. It has caught the attention of policymakers worldwide and similar funds have been introduced in Africa, Asia and other parts of Latin America. In Latin America and the Caribbean alone, 16 Social Investment Funds have been created to date (Glaessner et al, 1994).

For the ex-post evaluation of the welfare impact of the SIF, pre-intervention baseline data have been collected in selected rural provinces. In June 1993, a baseline survey was fielded in five provinces in the Chaco area.² In October and November 1993, this survey was extended to 16 other provinces in selected rural areas throughout Bolivia (hereafter named Resto Rural; see the Appendix for a map of the provinces in the survey).³ The survey gathered information on communities, facilities and individuals that are expected to benefit from projects financed by the SIF. The survey also collected information for control groups, which are not expected to receive SIF interventions. Interestingly, two different methods have been used to construct the control groups. In the Chaco region an experimental design has been used while in the Resto Rural the control group has been constructed by matching on observed characteristics.

This paper focuses on two issues that can be studied with pre-intervention data, and that are important for the impact evaluation. First, we look at the target population of the SIF and compare it with the total population in the Chaco and the Resto Rural. We investigate the effects of the institutional design of the SIF on its targeting. The demand driven approach of the SIF does not necessarily guarantee that the poorest in the society will benefit from the investments. Second, for the education component of the SIF we investigate the adequacy of the two evaluation designs that have been proposed: randomization and matched comparison. Baseline data allow us to compare control and treatment schools before the intervention, and hence to test for nonrandom selection of the control group and for a potential bias due to a selective control group. We find evidence of such a bias in the Resto Rural region. As a result, a simple comparison of treatment and control schools would lead to biased estimates of the welfare impact of the program for that region. This result is in

¹ As of July 31, 1995. Source: "Evaluacion de Medio Termino-Credito AIF 2532", FIS, September 1995.

² The provinces in the Chaco region are: O'Connor and Gran Chaco in the department of Tarija; Cordillera in the department of Santa Cruz; and Luis Calvo and Hernando Siles in the department of Chuquisaca.

³ The provinces in the Resto Rural region are: Bernardo Saaverdra, Camacho, Muñecas and Franz Tamayo in the department of La Paz; Capinota, Tapacarí, Quillacollo and Arque in the department of Cochabamba; Saucarí, Cercado, Carangas, Sur Carangas and Nor Carangas in the department of Oruro; and Linares, Nor Chicas, Saavedra and Modesto Omiste in the department of Potosí.

line with the finding of LaLonde (1986), who reports poor performance of a matched-comparison design in the evaluation of a social program in the US.

We propose to use instrumental variable (IV) techniques to correct for the selectivity bias. Usually, the choice of an instrument is hard to justify. However, with the pre-intervention data we can test whether an instrument is valid. We find that among several candidates that could be used as instrumental variables the number of NGOs (nongovernmental organizations) is a valid instrument, because it has a significant effect on the selection into the treatment group, but it does not affect the relevant response variables. The IV estimator is less efficient than the estimator obtained from a randomized design. We study this loss of efficiency, and we estimate the sample size that would be needed to make up for this loss.

The paper is organized as follows. In section 2 we present the data and sample design. In section 3 we describe the SIF in greater detail and look at the targeting mechanisms used by the SIF. In section 4 we compare the evaluation designs for the education component. In section 5 we propose the instrumental variable method that eliminates the bias in the matched-comparison design, and we study the loss of efficiency that comes with this method. Section 6 contains our conclusions.

2. Data

In the impact evaluation three types of investment projects are considered: health care, education and sanitation. The data collection for the impact evaluation is based on surveys applied to the both the institutions that receive funding (schools and health centers) and the households and communities that benefit from the investments. Similar data was also collected from comparison institutions and households.

The household survey consists of three subsamples. The first is a random sample of all households in the Chaco and Resto Rural regions, the second is a sample of households that live near the schools in the treatment or control group for the education component, and the third consists of households, that will benefit from the water and sanitation component. The surveys can easily be merged. For example, each school has a unique code which is recorded in the household survey if a child attends that school. The surveys for the Chaco and the Resto Rural region differ slightly. The baseline data collected in the Resto Rural are more extensive since shortcomings that were discovered in the surveys in the Chaco, that were conducted first, were corrected. Sample sizes are given in table 1.

The health facility survey gathered information on the quality of the infrastructure, staffing and visits to the center. Since the SIF planned to intervene in all health centers in the Chaco and the Resto Rural, all were included in the survey. The survey distinguishes between health clinics at the sector, area and district levels. Sector health clinics are typically very small, providing basic health care. Area-level health clinics provide more sophisticated care and serve a larger geographical region. District-level health clinics are hospitals, the largest type of facility. The larger the health clinic, the more detailed the questionnaire that was administered. The questionnaires are, however, comparable and collected similar types of information.

The school survey used two questionnaires, one for the director and one for each professor separately. It gathered information on infrastructure, equipment, teaching methods, and dropout and repetition rates of students. For the Chaco region, the sampling

frame consisted of all primary and secondary schools which qualified for SIF interventions. Eligibility was based on a school quality index, which was constructed using a weighted average of six infrastructure and equipment indicators. Schools were ranked according to this index, where a higher value reflects fewer resources. Only schools with an index above some cutoff value were eligible. The same procedure was used in the Resto Rural.

The community survey collected data from community leaders on a range of topics, from the quality of the infrastructure and distance to facilities to the presence of local organizations.

The household survey gathered information on each household member's education, health status and consumption.

Table 1. *Sample sizes of baseline survey*

		Chaco	Resto Rural
<i>Health centers</i>	Sector	82	119
	Area	16	24
	District	3	4
Total Health Centers		101	147
<i>Schools</i>	Treatment	35	37
	Control	37	33
Total Schools		72	70
<i>Households</i>	Random sample	2029	2138
	For education component	995	902
	For sanitation component	666	569
Total Households		3690	3609

3. The targeting of SIF interventions

The SIF has traditionally funded but not executed project proposals received from the private, public and not-for-profit sector. The SIF is a demand-driven institution because it does not initiate projects, but responds to outside initiatives by providing co-financing for investments in infrastructure, equipment and training. The co-financing provided by the SIF has generally accounted for approximately 80% of project costs, with the remaining 20% to be provided by the requesting institution. Regional SIF offices provide assistance to communities in preparing proposals and the decision on whether to fund a project is taken at the SIF central offices in La Paz. The final outcome thus depends both on the preferences and capabilities of the local communities, in particular the local authorities and local NGO's with respect to preparing and co-financing projects, as well as on centrally defined targets (see also Newman, Grosh and Jorgensen 1992). As a result, central objectives, such as targeting the poor, may not always be reflected in the final outcome of the program, particularly since the project approval process has tended to favor more well-organized

groups with access to counterpart financing which are not usually found in the poorer areas⁴. This is an inherent conflict of demand driven projects.

An analysis of the SIF's targeting mechanisms using the available pre-intervention data can not deal with behavioral responses that may result from SIF interventions. Baseline data can only be used to characterize households which are using the facilities in which the SIF is planning to invest. Of course, changes in household behavior as a result of changes in the supply of public services may be an important factor in determining the net impact of the project (Jimenez 1995). With the current data, however, we cannot deal with these effects.

Several aspects of the incidence of SIF interventions are presented in Figure 1. The graph in the left top corner is a nonparametric estimate⁵ of the density function of log per capita consumption. The estimate is based on the random sample of all households. The graph shows the distribution of consumption in the population and can be used as a guide for assessing the levels and concentration of poverty in the population. The graph in the top right corner presents a nonparametric regression estimate of the probability that an individual visited a government health clinic in the month before the survey date as a function of log per capita household consumption. This estimate is based on the random sample of all households. The households that have a high probability of visiting a government health clinic are more likely to benefit from the SIF investments. The lower left graph shows a nonparametric regression estimate of a function that is proportional to the conditional probability that a child in a household attends a school that will receive SIF funding, given the log consumption per head in the household. The estimate is based on the equality

$$\Pr(SIF\ school|\log(Cons / Head)) = \frac{f(\log(Cons / Head)|SIF\ school)}{f(\log(Cons / Head))} \Pr(SIF\ school)$$

The probability that a child in the household attends a school that will receive support from the SIF is not known, but from the random sample of all households and the sample of households near schools that are selected by the SIF, we can estimate the densities in the numerator and denominator of the first term on the right-hand side of the equation. The estimated ratio is depicted in the lower left-hand corner. The nonparametric regression estimate in the lower right-hand corner was obtained in the same way, be it that the event of interest is that the household lives in a community that is selected for an investment in sanitation. The graphs reveal that there is no relation between household welfare, as measured by log consumption per head, and benefits from SIF interventions in education . However, SIF investments in health facilities and basic sanitation benefit households that are relatively better-off.

⁴ With the introduction of the *Ley de Participacion Popular* in 1994 the preferences of the local populations may become more influential in the future. Under this law, a higher proportion of the government budget is allocated directly to the local communities. Municipal elections ensure accountability of the local leaders.

⁵ The nonparametric density and regression estimates are kernel smooths based on a Gaussian kernel. The bandwidth was chosen as suggested by Silverman (1986), p. 45.

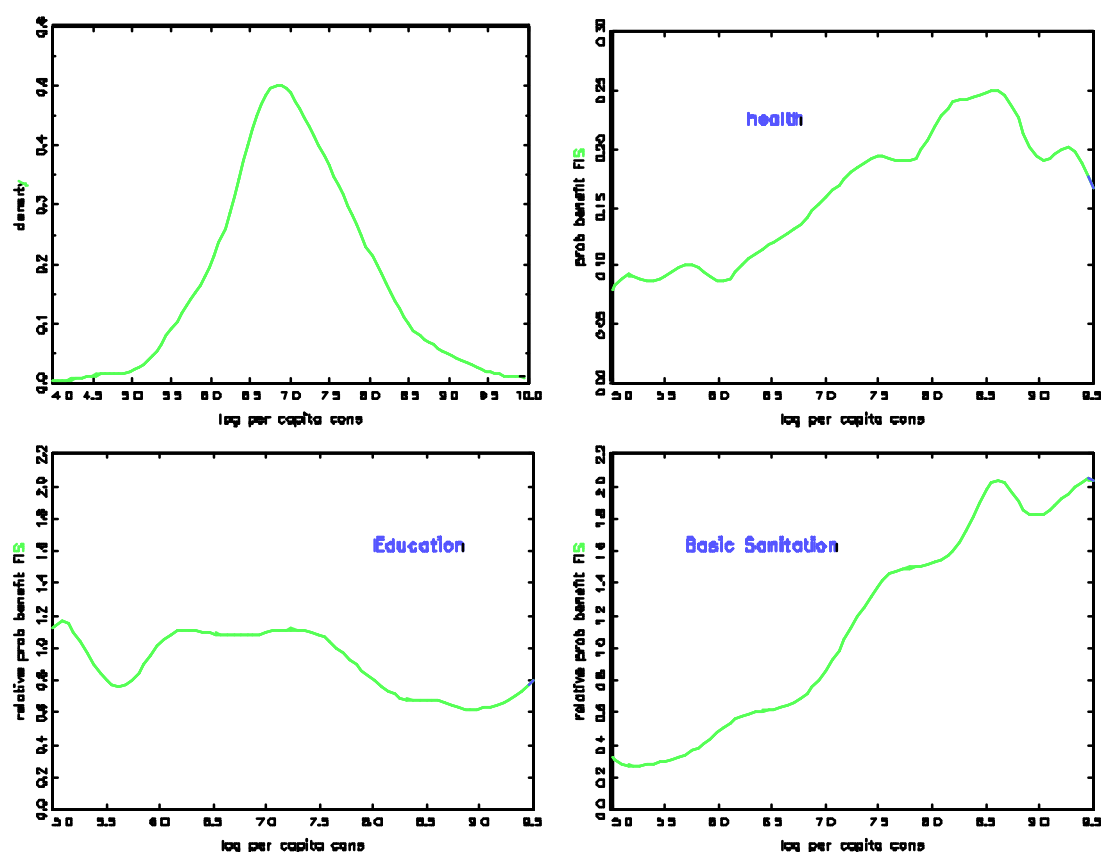


Figure 1. Targeting of SIF interventions: nonparametric estimates of the probability of benefit given household welfare.

District, area and sector health clinics are not the only providers of medical care. Private doctors and particularly traditional healers are extensively consulted for medical care as is indicated by Table 2, which was constructed to determine who will benefit from the SIF investments in health centers. To this end, the population has been divided into four groups ranging from richer to poorer depending on per capita consumption.

Table 2 presents results on the health status of future SIF beneficiaries and their use of health services prior to the SIF intervention. Health status is measured by asking the respondent whether he or she was in good health during the past month. Richer households tend to report a health condition that is worse than that of poorer households. However, this self-reported condition does not necessarily reflect the actual health status of the respondent as richer people may be prepared to admit to being in bad health more readily than poor people. Being in bad health is often associated with 'not being able to work' or 'seeking medical care' which is more readily affordable for the rich. What the data clearly show is that, if ill, rich households seek health care more frequently and go more often to government health clinics. Poor households seek medical care less frequently and visit traditional healers more often. If the poor visit government health clinics, they mostly go to small (sector) health clinics. The results suggest that, if the SIF wants to target investments in health facilities to poor communities, those investments should be concentrated in sector

and community-level health clinics. The results also clearly show a need for information and outreach programs to encourage poor households to seek medical care when ill and to visit public health care providers.

Table 2. *Health status, and actual health care consumption in the month preceding the survey(percentages)*

	Health status	Sought med. care	If sought medical care, went to:			If went to govt. clinic, type:		
			Govt. clinic	Private modern	Tradit.	Sect.	Area	Distr.
1-poor	84	44	44	10	47	48	33	19
2	83	55	58	10	32	42	35	23
3	80	58	59	20	21	38	34	28
4-rich	75	58	63	22	15	38	31	31
all	81	54	57	16	27	40	33	27

Source: SIF survey 1993

Note:-Health status is the percentage that reported to be healthy in the month previous to the survey.

-Modern private care includes private doctor and midwife. Traditional care includes traditional healers, neighbors, family and others.

-Quartiles correspond to the following per capita yearly consumption levels: quartile 1, up to 746 Bolivianos; quartile 2, 747-1224 Bolivianos; quartile 3, 1225-2158 Bolivianos; quartile 4, above 2158.

-sector clinics are small, community health centers; district hospitals are large, central hospitals

With respect to the sanitation and water investments, we find that households that will benefit from SIF investments in basic sanitation, have already better sanitation facilities than most of the rural population. For example, 47 percent of the targeted households have access to piped water while this only holds for 26 percent of the rural households in general. However, this is not necessarily inconsistent with SIF policy, since basic sanitation investments are most effectively made in areas that: (i) already have some infrastructure of this kind and (ii) are in populated areas so that the project is able to take advantage of economies of scale. Constructing these facilities in remote rural areas may lead to better targeting, but would end up being extremely costly relative to of the per capita benefits achieved.

The results of this section show that evaluation of the impact of the SIF investments may be problematic. The selective targeting of investments in health care and sanitation biases a direct comparison of beneficiaries and non-beneficiaries, because the latter were

already worse off before the intervention. An impact evaluation should take this selectivity into account⁶. A direct comparison may be possible for the education component, and in the next section we investigate, whether an unbiased estimate of the effect of SIF investments in education can be obtained.

4. Randomization and matched-comparison design

In the Chaco region eligible schools⁷ were allocated to the treatment and control group at random. Because of the backlog of existing commitments to fund school projects throughout the country, random selection of schools was not possible for the Resto Rural evaluation. Therefore, a control group was constructed by matching on a number of observable characteristics. The matching procedure consisted of two steps. First, using the 1992 census, cantons in which the treatment schools were located were matched to similar cantons with respect to: (i) population size and distribution by age and gender; (ii) literacy rate; (iii) infant mortality rate; and (iv) education level and language. Second, once the control cantons were identified, control schools were selected from the cantons to match the treatment schools with respect to the school quality index as developed for the Chaco region (see section 2).

As a consequence, two distinct evaluation designs are used in the two regions: a classical experimental design with randomized assignment in the Chaco and a matched-comparison design in the Resto Rural. In recent years, there has been a controversy over the validity of various evaluation designs. In particular, LaLonde (1986) has argued that matched-comparison designs can be severely biased, and that randomized assignment is the only design that can produce unbiased estimates of the effect of some intervention. In a reaction, Heckman and Hotz (1989) have argued that careful modeling of the selection effect can remove most of this bias. However, because of the uncertainties in this approach, it seems that a more secure basis for identification of the intervention effect is needed. In particular, a good model of the selection process can only be obtained if there is some exogenous variation in the selection into the treatment group. A variable that affects the selection into the treatment group but not the relevant response variable is a valid instrument, and if such an instrument is available it can be used to obtain an unbiased, be it less efficient, estimate of the intervention effect (Angrist 1989, Imbens and Angrist 1994). An instrument corresponds to an indirect experiment as opposed to a direct experiment due to randomization. Indirect experiments may be the only available evaluation design in many instances, because, as in the Resto Rural, random assignment is politically unfeasible.

In this section we check whether the assignment in the Chaco was indeed random, and we test whether the matching in the Resto Rural was selective and gives a biased estimate of the effect of the SIF. As the SIF co-finances investments in schools, it is natural to take the school as the unit of interest in the evaluation. The goal of the investment is to improve the quality of education, and hence we need variables that measure this quality. An

⁶ Under the assumption that changes that occur in the time between the baseline survey and the date of measurement of the response variables affect all relevant units, *i.e.* facilities, regions and communities, in the same way, we can estimate the impact of the SIF with a difference-in-differences estimator, as in Ashenfelter (1978).

⁷ Note that the restriction to eligible schools implies that the effects of the SIF can not be generalized to all schools.

obvious choice would be the (average) score(s) on a standardized test of the achievements of the pupils, but unfortunately such a test could not be administered. For that reason we use two indirect measurements of education quality: the fraction of the pupils that duplicates a class in the year of the survey, the duplication rate, and the fraction of students that drops out of school in the same year, the drop-out rate. It should be stressed that both variables are indirect measures of school quality, and in particular a high duplication rate may result from either high standards or low-quality education. The quality of education depends on the resources of the school, and the SIF aims at improving these resources. Hence, we compare the schools also on the available resources. In the sequel we use both the indirect measures and an indicator of the resources of the school as outcome variables. The results of the comparison are reported in Table 3.

Table 3a. *Descriptive statistics and comparison of treatment and control schools: Chaco*

	Treatment		Control		P-value of test of equal	
	N	Mean	N	Mean	Mean	Distribution
<i>Response var.</i>						
Duplication rate	35	.13	36	.13	.95	.86
Drop-out rate	35	.13	36	.09	.20	.51
<i>School resources</i>						
Blackboards per classroom	35	.12	37	.07	.43	1.00
Desks per student	35	.60	36	.44	.14	.06
Students per classroom	35	24.3	36	22.1	.36	.12
Books per student	27	.48	28	.31	.15	.09
Students per professor	35	19.9	36	20.0	.97	.90
Fraction profs with title	35	.37	37	.28	.34	.33
<i>Characteristics of students</i>						
Log per capita consump. of household	31	7.05	34	7.03	.81	.98
Education of mother (years)	31	2.49	34	2.05	.32	.37
Education of father (years)	31	3.49	33	2.71	.15	.22
<i>Community characteristics</i>						
Number of NGOs	28	.54	28	.14	.08	.44
Knowledge of SIF	28	.46	28	.46	1.00	1.00
Population size	28	454.1	28	468.4	.93	.90
Distance to main road (km)	28	17.8	28	13.3	.40	.44

Table 3b. *Descriptive statistics and comparison of treatment and control schools: Resto Rural*

	Treatment		Control		P-value of test of equal	
	N	Mean	N	Mean	Mean	Distribution
<i>Response var.</i>						
Duplication rate	24	.08	21	.08	.92	.98
Drop-out rate	24	.13	21	.08	.21	.62
<i>School resources</i>						
Blackboards per classroom	37	.53	33	.22	.00	.02
Desks per student	37	.72	32	.45	.03	.04
Students per classroom	37	24.1	33	24.5	.81	.56
Books per student	37	.38	33	.28	.27	.66
Students per professor	32	23.2	27	23.3	.97	.83
Fraction profs with title	32	.55	27	.57	.86	.88
<i>Characteristics of students</i>						
Log per capita consump. of household	32	6.57	32	6.61	.80	.53
Education of mother	32	1.35	32	1.43	.80	.94
Education of father	32	3.85	31	3.75	.86	.95
<i>Community characteristics</i>						
Number of NGOs	33	1.67	29	.72	.00	.01
Knowledge of SIF	33	.79	29	.45	.01	.03
Population size	33	293.4	29	344.7	.52	.10
Distance to main road (km)	33	13.6	29	9.56	.33	.69

Note:-Household consumption is in Bolivianos. In 1993 1 US\$=4.31 Bolivianos.
The test of equal means is the t-test; the test of equal distributions is the Kolmogorov-Smirnov test.

Table 3a shows that in the Chaco region the selection for the SIF was indeed random. There are no significant differences between the treatment and control schools either in a comparison of the response variables, or in a comparison of the school, student or community characteristics. In contrast, in the Resto Rural the matching of treatment schools with control schools on observable characteristics has not eliminated all differences between the schools. Although there are no significant differences for the response variables, the selected schools have significantly more resources. Moreover they are located in communities with a larger number of NGOs and greater knowledge of the SIF.

Table 3 contains univariate comparisons. As a further check on treatment assignment we estimated a probit that relates the probability of being a control school to the resources of the school and the characteristics of the student population. The results are in Table 4.

Table 4. *Probit estimate of probability of selection for control group*

	Chaco region		Resto Rural	
	Coefficient	t-value	Coefficient	t-value
<i>School resources</i>				
Blackboards per classroom	-.36	-.53	-1.56	-3.17
Desks per student	-.80	-1.98	-1.17	-2.27
Students per classroom	-.023	-1.10	-.042	-1.43
Books per student	-1.10	-2.30	.29	.45
Students per professor	-.026	-.73	.0011	.026
Fraction profs with title	-.62	-1.46	.12	.23
Dummy for missing school data	0.18	0.42	-0.097	-0.082
<i>Characteristics of students</i>				
Log per capita consump. of household	-.29	-.66	.32	.73
Education of mother	-.081	-.42	.29	1.17
Education of father	-.12	-.78	.036	.30
Dummy for missing student data	2.40	.78	-2.27	-.78
Constant	2.05	1.96	1.72	1.45
LR-test of significance (p-value)	14.1 (.23)	-	20.0 (.046)	-
Number of observations	71	-	69	-

The probit estimates confirm the results of the univariate comparisons in Table 3, be it that there is weak evidence that even in the Chaco region schools with more resources have a higher probability of being selected for the SIF. However, the likelihood ratio test does not reject the hypothesis of random selection. That hypothesis is rejected for the Resto Rural, and again because better-off schools have a higher probability of selection. Note that the dummy for missing student data is not significantly different from 0, so that the selectivity is indeed due to non-random selection and not to non-randomly missing observations.

Because the assignment to the SIF seems to be related to the resources of the school, and the resources affect the response variables we estimated a linear regression, in which the duplication rate and the dropout rate are related to the resources, the student characteristics and an indicator of being a control school. The results are in Table 5.

Table 5a. OLS estimates regression of response variables on school/student characteristics and indicator of control group: Chaco region

	Duplication rate		Dropout rate		Desks per student	
	Coeff.	t-value	Coeff.	t-value	Coeff	t-value
<i>School resources</i>						
Blackboards per classroom	.026	.43	.032	.61		
Desks per student	-.055	-1.53	.035	1.14		
Students per classroom	.0022	1.23	.0020	1.28		
Books per student	.010	.24	.062	1.70		
Students per professor		.26	-.006	-2.59		
	.00075					
Fraction profs with title	.013	.33	-.021	-.62		
Dummy for missing school data	.019	.47	-.019	-.54		
<i>Characteristics of students</i>						
Log per capita consump. of household	.046	1.20	.040	1.24	-.19	-1.42
Education of mother	-.034	-1.89	-.013	-.86	.0037	.063
Education of father	.022	1.71	.0070	.642	.043	.94
Dummy for missing student data	-.32	-1.23	-.29	-1.28	.99	1.08
Control group	.0039	.123	-.014	-.53	-.13	-1.24
Constant	.076	.81	.19	2.37	.741	4.42
R^2	.21	-	.22	-	.10	-
No. of observations	71	-	71	-	71	-

Table 5b. OLS estimates regression of response variables on school/student characteristics and indicator of control group: Resto Rural

	Duplication rate		Dropout rate		Desks per student	
	Coeff.	t-value	Coeff.	t-value	Coeff	t-value
<i>School resources</i>						
Blackboards per classroom	-.014	-.38	-.056	-1.37		
Desks per student	-.067	-1.76	-.084	-2.02		
Students per classroom	.0014	.39	-.0010	-.28		
Books per student	.011	.18	-.060	-.91		
Students per professor		.15	-	-.12		
	.00052		.00046			
Fraction profs with title	.0025	.051	.084	1.58		
Dummy for missing school data	.020	.20	-.20	-1.84		
<i>Characteristics of students</i>						
Log per capita consump. of household	.018	.40	.027	.53	.063	.56
Education of mother	-.008	-.49	-.024	-1.31	.11	1.86
Education of father	-.001	-.12	.020	1.88	.041	1.09
Dummy for missing student data	-.15	-.49	-.37	-1.12	-.87	-1.18
Control group	-.026	-.78	-.094	-2.61	-.26	-2.30
Constant	.12	1.02	.54	4.21	.85	4.77
R^2	.28	-	.58	-	.26	-
No. of observations	44	-	44	-	69	

Besides the duplication and dropout rates, we also use an indicator of the resources of the school, the number of desks per student, as a response variable. For the Chaco region the indicator of the control group is not significantly different from 0 for all three response variables. Again this confirms the effectiveness of using random assignment to generate equivalent treatment and control groups. For the Resto Rural we find that before the intervention the control group had lower duplication and dropout rates and fewer desks per student. As noted before control schools have fewer resources, but have duplication and dropout rates that do not differ significantly from treatment schools. For that reason, we find that, controlling for differences in resources, schools that perform poorly with respect to the duplication and dropout rates are selected to receive SIF funding. The difference is significant for the dropout rate and the number of desks per student. Note that none of the indicators for missing observations has a coefficient that is significantly different from 0.

The results also give an indication of the quality of the response variables. The duplication rate is weakly correlated with the resources of the school. That correlation is stronger for the dropout rate. This confirms our earlier doubts on the use of the duplication rate as a response variable.

We conclude that the matched-comparison design for the Resto Rural is biased. Control schools have fewer resources, but make better use of their resources, which results in lower duplication and dropout rates. The resulting bias of the estimator of the SIF effect, based on a direct comparison between treatment and control schools can be derived from the linear regression equation that relates the response variable y to the vector of school resources x (we omit student characteristics, which do not differ between treatment and control schools, and are assumed to be the same before and after the intervention)

$$y_{stk}^v = \alpha_{st}^v + \beta' x_{stk}^v + \varepsilon_{stk}^v \quad t = 0,1 ; k = 1, \dots, K_s ; s, v = C, T$$

where k is the school, s indicates the assignment to the treatment or control group, v denotes the outcome with or without the SIF investment, and t is 0 if the data are pre-intervention. We introduce the superscript v to stress that the treatment effect is defined for the same group of schools, indexed by either C or T , but that we can only observe schools with $s = v$.

The true SIF effect can be defined either for the treatment or the control schools. If we choose the first option, the SIF effect is

$$(\alpha_{T1}^T - \alpha_{T1}^C) + \beta'(\bar{x}_{T1}^T - \bar{x}_{T1}^C)$$

Taking the average over the schools in the treatment and control groups, the post-intervention difference in the average response is

$$(\alpha_{T1}^T - \alpha_{C1}^C) + \beta'(\bar{x}_{T1}^T - \bar{x}_{C1}^C)$$

so that the bias of the post-intervention difference as an estimator of the SIF effect is equal to

$$(\alpha_{T1}^C - \alpha_{C1}^C) + \beta'(\bar{x}_{T1}^C - \bar{x}_{C1}^C)$$

The bias depends on the contrasts between the treatment and control schools for the counterfactual case that the treatment schools do not receive the SIF support.

Under the strong assumption that

$$\alpha_{T1}^C - \alpha_{C1}^C = \alpha_{T0}^C - \alpha_{C0}^C \quad \text{and} \quad \bar{x}_{T1}^C - \bar{x}_{C1}^C = \bar{x}_{T0}^C - \bar{x}_{C0}^C,$$

and the weak assumption that

$$\alpha_{T0}^C = \alpha_{T0}^T \quad \text{and} \quad \bar{x}_{T0}^C = \bar{x}_{T0}^T,$$

we find that the bias is equal to the difference between the pre-intervention average responses

$$(\alpha_{T0}^T - \alpha_{C0}^C) + \beta'(\bar{x}_{T0}^T - \bar{x}_{C0}^C)$$

so that the difference between the observed post- and pre-intervention averages gives an unbiased estimate of the SIF effect for the schools that are selected for the SIF. The estimator is the difference-in-differences estimator, as used by Ashenfelter (1978). Because there is a four year period between selection and evaluation, it is unlikely that the assumptions will hold.

From the estimates in Tables 3-5 we have

$$\alpha_{T0}^T < \alpha_{C0}^C \quad , \quad \beta_{\bar{x}_{T0}^T} > \beta_{\bar{x}_{C0}^C}$$

and the pre-intervention difference between the average responses is 0. Note that this does not imply that the post-intervention difference in the average response is an unbiased estimator of the SIF effect. The strong assumptions introduced above are needed to guarantee that this estimator is unbiased. We conclude that the fact that the difference between the pre-intervention average response of the treatment and control groups is 0 is neither a necessary nor a sufficient condition for the unbiasedness of the post-intervention difference as an estimator of the SIF effect for the schools that receive SIF support.

5. A quasi-experimental evaluation design

A number of methods have been suggested to deal with selective treatment assignment. The first approach is to model the treatment assignment. The estimates in Table 4 are a first step in this direction. However, as the results in Table 5 show, the bias is not induced by the observable variables in the regression for the treatment assignment, but by the unobservables. Hence, we should allow for a correlation between the error of the regression for the response variable and the error of the probit model for treatment assignment (Heckman, 1976). To apply this method we must specify the joint distribution of the unobservable errors. It is well-known that this method only gives sensible results, if there are variables that affect the treatment assignment, but not the response variable. Such variables are also essential for the quasi-experimental approach, which does not require arbitrary distributional assumptions, and for that reason is more robust.

The quasi-experimental approach starts from the observation that randomized assignment as used in a classical experiment induces exogenous variation in the intervention indicator, and this variation has no effect on the response variable. If we can find a variable that affects the treatment assignment, but not the response variable, we have exogenous variation that mimics the type of variation induced by randomization. A variable with these properties is called an instrumental variable, and the corresponding ‘experiment’ is referred to as a quasi- or indirect experiment. The instrumental variable estimator of the intervention effect is unbiased, but it is less accurate, *i.e.* it has a larger variance, than the estimator of the intervention effect that could be used if the treatment assignment had been random.

In this section we show that community characteristics provide valid instruments. Some of these affect the selection into the SIF, while they do not have an effect on the pre-intervention response variables. Hence, these instruments can be used in a quasi-experimental evaluation of the SIF. We also study the efficiency of this design. We show that the relative efficiency, *i.e.* relative to a randomized design, is independent of the true treatment effect, and hence can be estimated using baseline data only. This allows us to

determine the sample size that will compensate for the loss of efficiency due to the quasi-experimental design.

The comparison in Table 3 suggests that for the Resto Rural the number of NGOs and knowledge of the SIF by community leaders have a significantly positive effect on the selection into the SIF. This is confirmed by a linear regression, the first step in a 2SLS estimation procedure for the SIF effect, of the indicator of selection into the program on the exogenous variables and the set of potential instruments of Table 6. The number of NGOs and knowledge of the SIF have a significantly positive effect in this linear probability model with coefficients .23 and .19, respectively. This could be expected, because of the fact that NGOs often act as subcontractors for the implementation of projects and the role that local leaders have in the selection process.

The variables are valid instruments, if they have no effect on the response variables. The evidence for the Resto Rural is in Table 6.

Table 6. OLS estimates regression of response variables on school/student characteristics and potential instruments: Resto Rural

	Duplication rate		Dropout rate		Desks per student	
	Coeff.	t-value	Coeff.	t-value	Coeff	t-value
<i>School resources</i>						
Blackboards per classroom	-.013	-.32	-.053	-1.07		
Desks per student	-.036	-.88	-.078	-1.59		
Students per classroom	.0015	.45	-.0012	-.30		
Books per student	.0037	.060	-.065	-.88		
Students per professor	.0005	.14	.00006	.013		
	2					
Fraction profs with title	-.043	-.82	.082	1.32		
Dummy for missing school data	.033	.32	-.21	-1.68		
<i>Characteristics of students</i>						
Log per capita consump. of household	.040	.81	.030	.51	.045	.37
Education of mother	-.021	-1.11	-.022	-1.00	.14	1.91
Education of father	.0007	.065	.019	1.44	.032	.80
	2					
Dummy for missing student data	-.27	-.83	-.38	-1.00	-.81	-1.00
<i>Potential instruments</i>						
Knowledge of SIF	-.035	-.84	-.0026	-.053	.067	.44
Number of NGOs	0.022	1.21	-.011	-.49	.0024	.031
Population (/1000)	0.038	.70	.010	.16	-.26	-1.20
Distance to main road	-.0010	-.80	-.00034	-.23	.0017	0.41
Dummy for missing data instr.	-.097	-1.44	-.00064	-.008	.20	.90

Control group	-.021	-.51	-.11	-2.17	-.22	-1.54
Constant	.19	1.52	.55	3.77	.71	3.14
F-test instruments (p-value)	1.22	(.32)	.10	(.99)	.60	(.70)
R^2	.42	-	.59	-	.29	-
No. of observations	44	-	44	-	69	-

Because the regression coefficients of the potential instruments are not significantly different from 0, we conclude that they are valid instruments.

The instrumental variables allow us to obtain an unbiased estimate of the effect of the SIF intervention from the post-intervention data from the regression equation

$$y_{s1k} = \alpha_{01} + \alpha_{11}d_{sk} + x'_{s1k}\beta + \varepsilon_{s1k} \quad k = 1, \dots, K_s; s = C, T$$

with $d_{sk} = 1$ for the schools that receive SIF funding. If X is the matrix with the observations on the independent variables, including the constant, but except the SIF indicator, and Z is the matrix with the observations on the instrumental variables, and we define

$$M = I - X(X'X)^{-1}X'$$

with I the identity matrix, then the variances of the OLS and instrumental-variable estimators of α_{11} are given respectively by

$$\frac{\sigma^2}{d'Md}$$

and

$$\frac{\sigma^2}{d'MZ(Z'MZ)^{-1}Z'Md}$$

with d the vector of observations on the SIF indicator. The ratio of these variances does not depend on the variance of the disturbance of the post-intervention regression σ^2 , and hence can be computed with pre-intervention data.

For the Resto Rural we find that the ratio of the standard errors of the estimates of the SIF effect is 3.97. Hence the standard error is four times larger than could have been obtained if the matched comparison had succeeded or if the SIF assignment had been random. As a consequence 242 schools instead of 61 are needed to estimate the SIF effect with the same precision as with randomized assignment. Because the assignment was random for the Chaco we find that the ratio is 11.20 for that region: using an instrumental-variable estimator with randomized assignment gives a very inaccurate estimate of the SIF effect.

6. Conclusion

We used pre-intervention data to study the targeting and selection into the SIF. For the health care and sanitation component we found that households that are better off will benefit more from the SIF investments. For these components the selectivity of the SIF complicates the impact evaluation.

For the education component an attempt was made to manipulate the selection into the SIF to facilitate the evaluation of its impact. For the Chaco region schools were assigned randomly to the SIF or to a control group. In the Resto Rural an attempt was made to mimic randomized assignment by matching treatment and control schools on observable characteristics. We found that this attempt was not successful. We proposed an alternative indirect procedure to evaluate the intervention in the Resto Rural. The pre-intervention data allowed us to verify that our proposal will produce an unbiased estimate of the SIF effect. We computed the loss of accuracy due to the indirect experiment and estimated the number of schools that will be needed to obtain an estimate with a precision that is comparable to that obtained from randomized assignment.

References

- Angrist, J. (1989), "Lifetime earnings and the Vietnam era draft lottery: evidence from social security administrative records", *American Economic Review*, vol. 80, pp. 313-36.
- Ashenfelter, O. (1978), "Estimating the effect of training programs on earnings", *Review of Economics and Statistics*, vol. 60, pp. 47-57.
- Heckman, J.J. and V.J. Hotz (1989), "Choosing among alternative nonexperimental methods for estimating the impact of social programs: The case of manpower training", *Journal of the American Statistical Association*, vol. 84, pp. 862-80.
- Glaessner, P.J., K.W. Lee, A. Sant'Anna and J. De Antoine (1994) "Poverty Alleviation and Social Investment Funds - The Latin American Experience", World Bank Discussion Paper no 261, Washington DC.
- Imbens, G. and J. Angrist (1994), "Identification and estimation of local average treatment effects", *Econometrica*, vol. 62, pp. 467-76.
- LaLonde, R.J. (1986) "Evaluating the econometric evaluations of training programs with experimental data", *American Economic Review*, vol. 76, pp. 604-20.

Jimenez, E. (1995), "Human and physical infrastructure: Public investments and pricing in developing countries", *Handbook of Development Economics*, vol. 3, ch. 47, North-Holland, Amsterdam.

Newman, J.L., M. Grosh and S. Jorgensen (1992) , "Demand -driven funds: Managing their conflicts", in: *Bolivia's Answer to Poverty, Economics Crisis and Adjustment*, World Bank Regional and Sectorial Studies, Ch 7, World Bank, Washington D.C.

Silverman, B.W. (1986), *Density estimation for statistics and data analysis*, Chapman&Hall, London.

Appendix

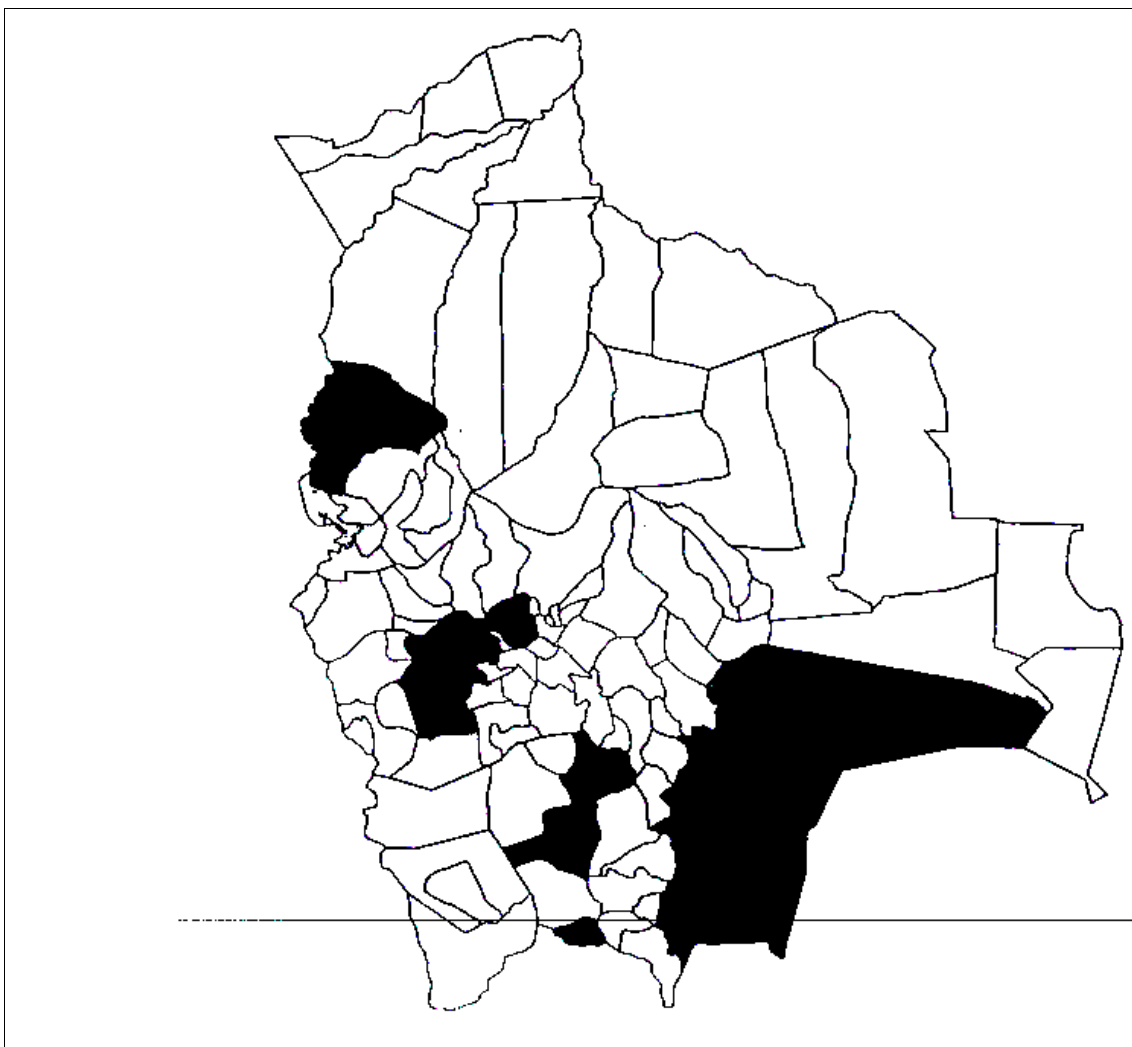


Figure A1 Selected provinces for SIF evaluation survey