

[1em]1em1em

Essays on Ex ante Evaluations of Cash Transfer Programs

Ranjeeta Alice Thomas

PhD Thesis

University of York

Department of Economics and Related Studies

September 2011

Abstract

This thesis comprises three essays on *ex ante* evaluations of cash transfer programs.

Chapter 2 uses baseline data from the randomized experiment of the conditional cash transfer program - Red de Protección Social (RPS), Nicaragua to forecast the impact on school enrolment and compares results to those of the experimental evaluation. Reduced form estimation of a behavioural model forms the basis of the evaluation. A Klein and Spady semi-parametric single index model is used to predict unobserved outcomes under the treatment. The sample consists of children aged 7-13 who have not completed grade 4 of primary school. The evaluation shows that the *ex ante* approach closely matches the experimental outcomes in most cases.

Chapter 3 extends the behavioural model of chapter 3 to include a health production function and analyses the impact of the program on health care utilization outcomes for infants. This chapter also uses baseline data from the randomized experiment, and applies the semi-parametric single index model to predict the *ex ante* impact. It validates the model with the results of the experiment and then simulates two alternate policy scenarios. The model performs well in predicting the health related outcomes and shows different results for the two sets of policy scenarios.

Chapter 4 also uses the data from Nicaragua's RPS social protection program to forecast the program's impact on quantiles of WAZ for children below 5 years. It applies the reduced form behavioural model of child health defined in chapter 3 to facilitate the empirical strategy. The estimation compares two semiparametric approaches - a quantile regression approach and a distributional regressions based approach with a nonparametric approach to estimate the unobserved unconditional counterfactual distribution under the program. In all cases the estimated effects are compared with those of the randomized experiment. The models perform well in forecasting the distributional impacts of the cash transfer of WAZ. Like in the experiment, the predicted outcomes show the greatest impact at the lowest quantiles of the distribution.

Contents

Abstract	1
Acknowledgements	6
Declaration	7
1 Introduction	8
2 Using a Semiparametric Estimator to Forecast Education Outcomes in Nicaragua’s Red de Protección Social	13
2.1 Introduction	13
2.2 Red de Protección Social	15
2.3 Economic Framework	18
2.4 Empirical Specification	23
2.4.1 Estimating School Costs	23
2.4.2 Estimating Counterfactual Outcomes	25
2.5 Data and Variables	29
2.5.1 Variables	29
2.6 Results	30
2.6.1 Estimating School Costs	30
2.6.2 Predicting Impacts	33
2.7 Conclusion	37
3 Conditional Cash Transfers to Improve Education and Health: An Ex ante Evaluation of Red de Protección Social, Nicaragua	38
3.1 Introduction	38
3.2 Background	40
3.2.1 CCTs and health services	40
3.2.2 Health care utilization and vaccinations	40
3.2.3 Red de Protección Social	41
3.3 Theoretical Framework	44

3.4	Empirical Specification	50
3.4.1	Estimating School Costs	51
3.4.2	Estimating Counterfactual Outcomes	52
3.5	Data and Variables	56
3.5.1	Dependent variables	57
3.5.2	Explanatory variables	58
3.6	Results	59
3.6.1	Estimating School Costs	59
3.6.2	Predicting Impact	62
3.7	Conclusion	68
4	Ex ante Forecasting of the Distributional Effects of a Conditional Cash Transfer Program on Child Health in Nicaragua	70
4.1	Introduction	70
4.2	Background	72
4.2.1	Malnutrition and CCTs	72
4.2.2	The RPS Program	73
4.3	Model and Empirical Specification	74
4.3.1	Model	74
4.3.2	Empirical framework	78
4.4	Data and Variables	84
4.5	Results	86
4.5.1	School costs	86
4.5.2	Ex post - distributional impact	87
4.5.3	Ex ante - distributional impact	90
4.6	Conclusion	95
5	Conclusion	97
	References	101

List of Tables

2.1	Estimating School Costs	31
2.2	Predicted Impact	36
3.1	Basic vaccination schedule for up-to-date vaccinations	58
3.2	Estimating School Costs	61
3.3	Predicted Impact	64
3.4	Simulating Counterfactual Policy Scenarios	66
4.1	Estimates of the model for School Costs	88
4.2	Estimated Policy Effects	89
4.3	Estimated Policy Effects - All Quantiles	93

List of Figures

2.1	RPS Eligibility and Requirements. Source: Maluccio and Flores 2005	17
2.2	RPS Transfers. Source: Maluccio and Flores 2005	18
2.3	Variation	21
2.4	Comparing Observed and Predicted Outcomes	34
2.5	Trimming	35
3.1	RPS Eligibility and Requirements. Source: Maluccio and Flores 2005	43
3.2	RPS Transfers. Source: Maluccio and Flores 2005	44
3.3	Variation	49
3.4	Comparing Observed and Predicted Outcomes	62
3.5	Trimming	63
4.1	Variation	76
4.2	2002 Quantile functions	80
4.3	Ex-post Quantile Treatment Effects	90
4.4	Ex ante Quantile Treatment Effects	91

Acknowledgements

I thank my supervisor Prof. Andrew Jones and the members of my Thesis Advisory Panel - Prof. Nigel Rice and Prof. Peter Smith for their support. I also thank Dr. Tim Beatty for his comments and encouragement while on the advisory panel for the first two years of my PhD. My gratitude to Dr. Lyn Squire for the inspiration and support. My deepest gratitude to my mother for without whose support none of this would have been possible.

Declaration

All chapters in this thesis are single-authored.

Chapter 3 is currently under review at the *Journal of Health Economics*. During the Eighth World Congress of the International Health Economics Association (iHEA), held in Toronto, Canada in July 2011, this version of the paper was awarded First prize in the iHEA student competition.

This thesis uses data made available by the International Food Policy Research Institute (IFPRI). All errors remain my responsibility.

Chapter 1

Introduction

This thesis comprises three essays on forecasting the impact of social programs on education and health outcomes. It tests a recently proposed approach by Todd & Wolpin (Forthcoming) and uses as a benchmark results from a randomized experimental evaluation of the social program to validate the results of the evaluation.

The motivation for testing the approach of Todd & Wolpin (Forthcoming) lies in increasing the applications of *ex ante* evaluations particularly in resource constrained settings in developing countries, not as a substitute for *ex post* experimental evaluations but as a useful complement to observe changes that can be expected. *Ex ante* evaluations differ from *ex-post* applications in that they focus on simulating the effects of hypothetical/yet-to-be launched programs or large changes in the parameters of existing programs. They involve extrapolations from existing policy or variation in policy-relevant parameters. Typically, *ex-ante* evaluation requires knowledge of past and future values of relevant exogenous variables and if future changes are expected in parameters that have not previously changed then knowledge of the structure of the relevant system is necessary (Marschak 1953).

Ex ante evaluations offer several advantages especially in development settings. In resource constrained situations, *ex ante* evaluations of varying policy or program intensity can help arrive at an optimal program design, without which multiple social experiments would need to be carried out, involving significant financial investments. *Ex post* evaluations are critical to assessing aid effectiveness, *ex ante* evaluations further enhance its effectiveness by ensuring investments are not made in programs that are likely to have no or limited impact. They can provide essential information on the target population most likely to benefit from the

program while at the same time providing estimates of impact for changes in parameters of ongoing programs. While *ex post* evaluations are critical in development, *ex ante* evaluations are a useful complement offering scope for furthering aid effectiveness.

Ex ante evaluations like *ex post* methods depend on household surveys. However, in contrast to *ex post* methods, *ex ante* evaluations involve generating a counterfactual representative of the population likely to benefit from a hypothetical program. Changes in individual behaviour under the counterfactual in response to policy change are then simulated to measure causal effects. Such evaluations are then marginal as they capture the difference between the status quo and a change in policy. These changes may or may not include behavioural responses (Bourguignon & Ferreira 2003). The approaches to *ex ante* evaluations differ in their use of microsimulation models or behavioural models that depend on structural or reduced form estimations.

This thesis focuses on the last of these approaches. This approach varies from the other two in that *ex ante* impact estimation is conducted without structural estimation and using non-specific functional forms. The application is a modification on earlier work by Ichimura and Taber (2000) who provide a general set of conditions for reduced form estimation under which new programs can be evaluated using relatively weak structures and full estimation of behavioural models is not required to identify program impacts. In their paper Todd and Wolpin consider a variety of policies including wage, income and school subsidy programs. Identification exploits variation in the policy variable to predict program impacts. In cases where no variation exists in the policy variable their approach depends on variation in other model variables that are indirectly related to the program. The key contribution of their paper is in placing *ex ante* evaluations in the potential outcomes framework. Program impacts are measured by a matching estimator that uses only control group (untreated) data and matches untreated individuals with other untreated individuals based on functions of observable variables generated by the model. Using simpler economic models and in most cases a fully nonparametric approach the authors show identification of conditions when such approaches can be applied. Their motivation is to provide easily applicable methods to forecast the impact of a wide range of social programs and interventions. This thesis tests their suggested approach by applying the method to a conditional cash transfer program in Nicaragua. This

program is particularly suited to this exercise as it was implemented as a randomized social experiment and evaluated in full. It also provides data on a range of outcomes covering both education and health - both binary and continuous.

The thesis is structured as a collection of three papers. Although they are presented as three separate papers, they all have as a common theme the approach of Todd and Wolpin, also all three papers use data from the randomized experiment in Nicaragua but use different samples depending on the outcome being estimated. They also have in common some estimation techniques and results which arise from the behavioural model and relate to the policy related variables.

Chapter 2 begins the three applications by estimating the outcome of the RPS program on school enrolment. The program provided a cash transfer conditional on school enrolment for all households in the treatment group with children aged 7-13 who had not completed grade 4 of primary school, the estimation sample consists of individuals who meet this criteria. The program also included a food security transfer. This chapter is essentially a direct application of the idea in Todd and Wolpin's paper but extends the behavioural model to include time costs. The behavioural model is a household maximisation problem that models school choices subject to a time constraint for the transfer eligible child and a household budget constraint. The time constraint models a choice between the time a child spends working and the time a child spends in school. The optimal choice of school enrolment is estimated as a function of observable covariates generated by the behavioural model. In addition assumptions are made about unobservable heterogeneity that require conditioning on a further set of covariates. In essence this approach relies on "selection on observables". The estimation of the binary outcome is implemented using a semi-parametric single index model. The paper uses data from the baseline survey of the experiment. The variation in the two policy variables - full income of the household (which is proxied by 'consumption' and school costs are used to forecast the outcomes. The school costs are however observed in the survey only for those enrolled in school at the time of the survey and are estimated for the other children in sample. The *ex ante* impact is the difference between the estimated outcomes under treatment and the observed outcomes pre-treatment. The results are compared with the one year effects of the program from the double difference estimates of the experiment.

Chapter 3 extends the behavioural model and analysis in chapter 2. It introduces a health production function in the behavioural model and derives the reduced form for the optimal choice of health care utilisation. Two preventive care outcomes for children under 5 years are considered in this paper - health checks and full coverage of vaccinations. The estimation procedure follows the same approach as chapter 2. The estimated school costs for the sample of older children are also incorporated into the model exploiting the variation in this variable along with full income (proxied by consumption). Given that chapter 3 is an extension of chapter 2 but written up as a separate paper, several segments of the two chapters are identical. For instance, the section on school cost estimation is the same in both papers as they are estimated only once for all children in the baseline survey, aged 7-13 years who have not completed grade 4 of primary school. The health care utilization outcomes used in this chapter are also binary measures and hence the empirical strategy section describing the semi-parametric estimator is also similar. The analysis in this paper is however extended beyond a comparison of experimental outcomes, to simulate counterfactual policy scenarios. The analysis enables comparing different trade-offs families make between education of older children and preventive health care for infants under different program designs. Two alternate policy specifications are simulated; reducing the amount of the transfer by 25% while maintaining the conditionalities and an unconditional transfer equal to the food security component of the program.

Chapter 4 is a further extension of the *ex ante* approach applied in chapters 2 and 3. The objective in this paper is to forecast the impact of the program on the entire distribution of the outcome of interest. Heckman, Smith & Clements (1997) highlight the importance of considering more than just the average impact of a program and exploring for heterogeneity. This chapter tries to capture this aspect of program evaluation in an *ex ante* scenario. The underlying motivation is the same reduced form approach of Todd and Wolpin and thus relies on the same behavioural model as chapter 3 but is markedly different in its estimation strategy to recover the unobserved distribution under treatment i.e. the impacts at different quantiles of the outcome distribution. The underlying behavioural model is therefore not presented again in this chapter. The outcome of interest here is the weight-for-age Z score (WAZ) for children under 5 years. While the age group is the same as chapter 3 the sample is different in two ways; data for only the randomized out control group is used and the data is from the 2 year follow-up of the program (i.e. 2002) rather than the baseline data. The reason for this

approach is that the outcome in the experiment was measured only at the baseline (2000) and after two years (2002). To facilitate comparison with the experimental outcomes data from the 2002 control group is used. Once again, as in the earlier chapters estimation of school costs for program eligible children is necessary, but for the sample of children in the 2002 survey. Estimation of the counterfactual distribution is carried out using two semi-parametric approaches and a non-parametric approach. The parameters of interest in this chapter are the “quantile treatment effects” as opposed to the average effect of the earlier two chapters. The quantile treatment effect is the difference between the estimated counterfactual distribution and the observed distribution without treatment. Once again results are compared with the treatment effects estimated from the experiment.

While the data used in this thesis is from a conditional cash transfer program, the methods applied are applicable to any social program that channels its impact through a change in budget constraints. This thesis is a first step towards testing the recommendations of Todd and Wolpin of applying simple behavioural models and relaxing assumptions of functional form to forecast outcomes of such programs. Each of the chapters builds on the preceding one either to explore alternative outcomes or simulations or to extend the analysis beyond the average impact. But in all cases it tries to use the randomized experiment as the benchmark against which to compare the results of the *ex ante* exercise. The thesis explores the feasibility of the assumptions made under this approach to *ex ante* evaluations for different objectives and also tries to assess the limitations of such an approach. These are discussed in summary in the conclusion chapter of this thesis (chapter 5).

Chapter 2

Using a Semiparametric Estimator to Forecast Education Outcomes in Nicaragua's Red de Protección Social

2.1 Introduction

The last few decades has seen billions of dollars channelled to developing countries as international aid. Despite this impetus these regions continue to remain amongst the poorest in the world with some of the worst indicators of poverty and health. Improving the critical link between aid and outcomes requires ensuring resources are channelled to where they are likely to have the greatest impact (White 2006). Research on this link involves the evaluation of development programs to measure their impact. Most of this research has focussed on *ex post* evaluations of programs by either randomized-experimental allocation of the intervention or using observational approaches such as difference-in-differences, matching methods or regression discontinuity. On the other hand, applications of forecasting in economics have been widely applied in estimating demand and predicting impacts of macro-economic policies but there are comparatively few applications in evaluating social programs. In development settings with constraints on resources, *ex ante* evaluations are particularly useful in making informed decisions for extending the target population of an existing program. They also facilitate optimal usage of limited resources by ensuring governments make financial investments in programs that are likely to have a positive impact. These evaluations are useful in considering implementation of new programs and serve as complements to future *ex post* evaluations.

Leading examples of *ex ante* evaluations of a social program are Todd & Wolpin (2006) and Attanasio et al. (2005), who evaluate the impact of Mexico's Progresa Conditional Cash Transfer program by structurally estimating the parameters of a behavioural model that specifies the interactions of the program. In contrast to the structural estimation approach in a recent simpler reformulation, Todd & Wolpin (Forthcoming) build on the work by Ichimura & Taber (2000) and illustrate the use of reduced form estimation of behavioural models in evaluating social programs without specification of functional forms. The authors illustrate situations in which a non-parametric estimation strategy based on a behavioural model can be used to estimate *ex ante* impacts. This reduced form *ex ante* approach differs from *ex post* evaluations in the way it uses the traditional potential outcomes framework in that the data are observed for only the untreated population. In this case the counterfactual to be estimated are the outcomes for the population when treated rather than for the controls. Program impacts using the behavioural model reduced form (BMRF) approach are estimated from an underlying economic model and use variation in the variable through which the policy instrument operates for model identification.

The objective of this paper is to apply the reduced form estimation approach to a conditional cash transfer program (CCT) and to compare the predicted outcomes with results from a randomized experiment. It is based on an economic model of household consumption and uses data from the experimental evaluation of Nicaragua's Red de Protección Social (RPS), a CCT program for rural households in Nicaragua. The program aims at improving school enrolment and attendance of children aged 7-13 years who have not completed grade 4 of primary school and health and nutritional status of children below 5 years by supplementing household income through the cash transfers. The cash transfers are conditional on a certain minimum school attendance by children of recipient households and attendance at health workshops by mothers. The analysis in this paper focuses on education outcomes.

The estimation strategy uses variation in the costs of schooling, full income of households along with several household characteristics to determine the impact of the program. The large number of covariates determining outcomes does not permit fully non-parametric estimation. To overcome this, a semiparametric single index model for binary outcomes is used to predict impact. This paper presents the economic model and estimates the impact on

school enrolment using data from the randomized experiment.

2.2 Red de Protección Social

Red de Protección Social (RPS) is based on the design of Mexico's PROGRESA program and is the first CCT to be implemented in a low-income country¹. The program was introduced in 2000 and targets reducing financial barriers to accessing education and health care in rural Nicaragua. In 1998 data from the Living Standards Measurement Survey (LSMS) indicated that 48% of the population in Nicaragua was poor and 75% of this population lived in rural areas (World Bank 1998). The program was implemented in two phases. The first phase was designed as a pilot randomized experimental evaluation in two districts, Madriz and Matagalpa, based on the level of poverty and capacity of these districts to implement the program. In both these regions 80% of the rural population was poor and of this population 50% were extremely poor (Maluccio & Flores 2005, IFPRI 2001a). Phase Two of the project extended the program for a further 3 years. This paper uses data from the pilot phase of the program generated by the randomized experiment. From the two regions selected 42 *comarcas* or administrative units were selected based on a marginality index to participate in the pilot.

The program consisted of two demand side components - the first focusing on food security, health and nutrition and the second on education. Each eligible household received a "food security transfer" in alternate months based on two conditions, attendance by mothers at health education workshops held every other month and children under age 5 being brought for scheduled preventive health check-ups. The demand side health initiatives were complemented by supply-side enhancements including training and payment to private health care providers to ensure the increased demand from the program was met. The food transfer was a fixed amount and did not depend on the size of the family. The education component of the program consisted of two cash transfer components to families with children aged 7-13 years who had not completed grade 4 of primary school, conditional on enrolment and

¹This section is based on the description of RPS provided by Maluccio & Flores (2005) in the impact evaluation report of the randomized experiment

regular attendance by the children. The first was a lump-sum transfer provided as a fixed amount per family regardless of the number of eligible children, conditional on all eligible children enrolling in school. In addition a cash transfer for school supplies was provided for each eligible child, also conditional on enrolment. On the supply-side, incentives were provided to teachers to compensate for the additional monitoring and reporting required to ensure compliance with the program and the increase in class size from the enrolments. This supply side component was administered through the student who presented the teacher with the cash on going to school and is thus unlikely to affect school enrolment, which is the outcome evaluated in this paper. The evaluation in this paper and the *ex post* results are for the demand-side cash transfer components of the program. While Figure 2.1 presents a summary of the eligibility criteria and requirements for RPS.

The transfers target a reduction in the net price of schooling and food consumption to reduce short-term poverty while encouraging investments in human capital to eliminate long-term poverty. The amounts of the transfer include the Córdoba 2000 equivalent of US\$224 for food security and US\$112 for the educational component. The school supplies component for each eligible family was US\$21. Figure 2.2 presents a summary of the transfers. According to the *ex post* evaluation of RPS by Maluccio & Flores (2005) the food transfer was equivalent to 13% of annual household expenditures and families with one eligible child for the schooling components would receive an additional 8% of annual household expenditures. Beneficiaries that did not comply with the specific requirements associated with each component failed to receive the transfer for the particular component.

The randomized evaluation provides census data (for all eligible households and individuals) in the 42 selected *comarcas*, baseline data for the final selection of households based on the marginality index after assignment into treatment and control groups and follow-up data for the next two years. Since the objective of this paper is an *ex ante* evaluation, the focus is on data generated prior to the introduction of the intervention.

PROGRAM REQUIREMENT	HOUSEHOLD TYPE		
	Households with no targeted children (A)	Households with children aged 0-5 (B)	Households with children aged 7-13 who have not completed 4 th grade (C)
	(A)	(B)	(B) + (C)
Attend bimonthly health education workshops	✓	✓	✓
Bring children to prescheduled healthcare appointments Monthly (0-2 years) Bimonthly (2-5 years)		✓	✓
Adequate weight gain for children under 5 ^a		✓	✓
Enrollment in grades 1 to 4 of all targeted children in the household			✓
Regular attendance (85 percent, i.e., no more than 5 absences every two months without valid excuse) of all targeted children in the household			✓
Promotion at end of school year ^b			✓
Deliver teacher transfer to teacher			✓
Up-to-date vaccination for all children under 5 years ^b		✓	✓

a. The adequate weight gain requirement was discontinued in Phase 11, starting in 2003
b. Condition was not enforced.

Figure 2.1: RPS Eligibility and Requirements. Source: Maluccio and Flores 2005

Nicaraguan RPS eligibility and benefits in Phase 1		
	Program components	
	Food security, health and nutrition	Education
Eligibility		
Geographic targeting	All households	All households with children aged 7-13 who have not completed fourth grade of primary school
Demand-side benefits		
Monetary transfers	<i>Bono alimentario</i> (food security transfer) C\$2,880 per household per year (US\$224)	<i>Bono escolar</i> (school attendance transfer) C\$1,440 per household per year (US\$112) <i>Mochila escolar</i> (school supplies transfer) C\$275 per child beginning of school year (US\$21)
Supply-side benefits		
Services provided and monetary transfers	Health education workshops every 2 months Child growth and monitoring Monthly: Newborn to 2-year-olds Every 2 months: 2- to 5-year-olds Provision of antiparasite medicine, vitamins, and iron supplements Vaccinations (newborn to 5-year-olds)	<i>Bono a la oferta</i> (teacher transfer) C\$80 per child per year given to teacher/school (US\$6)

Figure 2.2: RPS Transfers. Source: Maluccio and Flores 2005

2.3 Economic Framework

The model uses the household production framework of Becker (1965). A household with multiple eligible children $i = 1 \dots n$, has utility U a function of C representing consumption, a vector S , with $S_i = 1$ a binary indicator of school enrolment of each child, and an indicator of gender g . The household maximisation problem is then:

$$\max_{C, S} U(C, S; g, \nu) \quad (2.1)$$

The time constraint for an eligible child can be written as:

$$T_i = T_{si} \cdot S_i + T_{wi}(1 - S_i) \quad (2.2)$$

where T_{si} is time spent in school and is assumed to be a fixed amount for all enrolled children, T_{wi} is time spent at work. This specification follows Todd and Wolpin and does not allow

for leisure in the time constraint as a substitute. However, a different form of time constraint that does not assume substitution is possible, that allows total time T_i to be divided between time spent on work, schooling and leisure. The reduced form in this case would be different to the one derived here.

The money budget constraint can be written as:

$$C + \sum_{i=1}^n \delta_i \cdot S_i = Y + \sum_{i=1}^n w_i \cdot T_{wi} (1 - S_i) \quad (2.3)$$

Where δ_i is the direct cost of schooling for child i . Primary schooling is free in Nicaragua and most children face no tuition fees, hence δ_i includes all other school related costs faced by families such as transport, uniforms, books and school meals. Y is household income net of the earnings of the program eligible children.

The constrained household maximisation problem is:

$$\max_{C,S} U(C, S; g, \nu) \quad (2.4)$$

The full income constraint combining both the time and money constraint is:

$$C + \sum_{i=1}^n [\delta_i + w_i \cdot T_{si}] S_i = Y + w_i \cdot \sum_{i=1}^n T_i = F \quad (2.5)$$

where F is full income of the household and the total price of schooling for all eligible children in the family ($\theta = \sum_{i=1}^n [\delta_i + w_i \cdot T_{si}]$) is the cost of schooling plus the shadow wage for the eligible children.

The optimal choice of schooling is $S^* = \Phi(F, \theta, n; g, \nu)$

The RPS program has two cash transfers - the first focuses on changing the price of schooling for eligible children conditional on enrolment and the second is a food transfer meant to boost consumption, nutrition and access to preventive health care conditional on mothers attending the health workshops (E_m). Where E_m represents a binary indicator of mother's attendance at compulsory health education workshops. Under the assumption of full compliance $E_m = 1$ in the post program scenario. The school transfer is implemented as two components (τ, ρ) to reduce the net price of schooling and substitute for any wages earned by children not enrolled in school due to employment. A decrease in the price of schooling is likely to encourage children to substitute away from labour market participation and increase school enrolment. The first component τ is provided for each eligible child in the family while ρ is a lump sum transfer irrespective of the number of eligible children. Both transfers are conditional on all eligible children enrolling in school. The household food transfer (μ) conditional on E_m is modelled as a direct income effect, raising the income level of the household and does not stipulate specific expenditure categories.

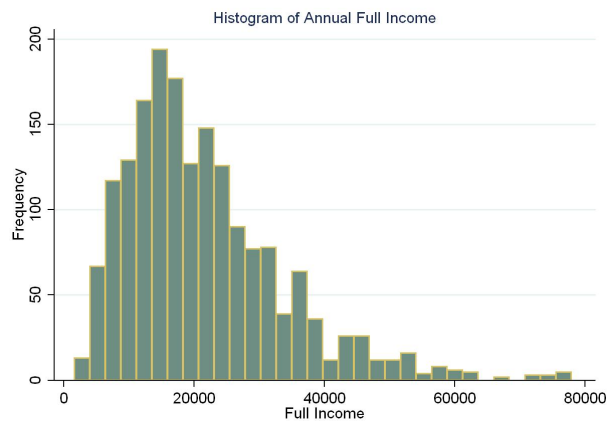
With the introduction of the subsidies $\mu.E^m$, $\tau.\sum_{i=1}^n S_i.S_p$ and $\rho.S_p$, where $S_p = 1$ if $\sum_{i=1}^n S_i = n$ ie. all eligible children enrol in school and $S_p = 0$ otherwise. The money budget constraint can be written as:

$$C + \sum_{i=1}^n \delta_i.S_i = Y + \sum_{i=1}^n w_i.T_{wi}(1 - S_i) + \mu.E^m + \sum_i^n \tau.S_i.S_p + \rho.S_p \quad (2.6)$$

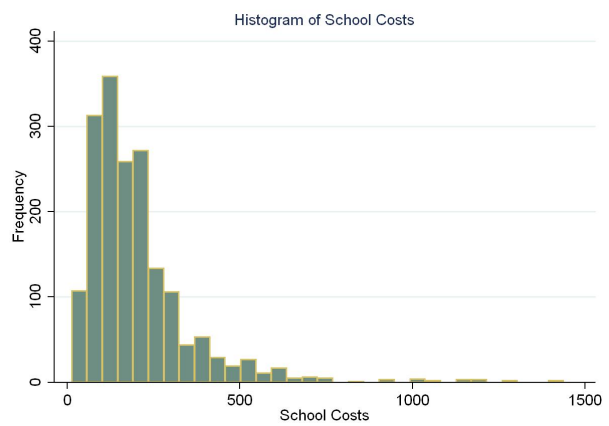
The full income constraint is then:

$$C - \rho.S_p - \mu.E^m + \sum_{i=1}^n (\delta_i + w_i.T_{si} - \tau.S_p) S_i = Y + \sum_{i=1}^n w_i.T_i = \tilde{F} \quad (2.7)$$

The new price of schooling under the subsidy program is $\tilde{\theta} = (\sum_{i=1}^n [\delta_i + w_i.T_s - \tau.S_p])$ and the cost of consumption is $C - \rho.S_p - \mu.E^m$. The optimal choice under the subsidies is $S^{**} = \Phi(\tilde{F}, \tilde{\theta}, n; g, \nu)$



(a) Household Consumption



(b) School Costs

Figure 2.3: Data Variation

Empirically this exploits two sources of variation in the data to compare untreated individuals with outcomes S^* with other untreated individuals with outcomes S^{**} - the first is school costs and the second is full income of the households at the baseline. As described earlier, primary education is free in Nicaragua and most families face no fees, the cost here includes other expenditure related to schooling which is exogenous in the sense that it is faced by all families when enrolling children irrespective of whether the tuition is free or not. In addition, the data do not provide a measure on income to be used as a proxy for full income, but has a measure of consumption. Figure 2.3(a) shows a histogram of consumption of families, with values ranging from c1,590 to c77,905. The second graph figure 2.3(b) shows the school costs used in the estimation range from c12 to c1438. In addition to variation in school costs and consumption, the level of the school grant also varies depending on the number of children in the household. The treatment effect is estimated by matching the treated and untreated groups on functions of observable characteristics. The key identification condition in this approach is that the program has an impact only through the budget constraint of the behavioural model (Todd & Wolpin Forthcoming), ensuring that the reduced form before the program is also the same after, except for a change in the magnitude of the variables resulting from the program. The approach also relies strongly on selection on observables to capture heterogeneity. As specified by Todd & Wolpin (Forthcoming) extending the approach to allow impact of the program to affect preferences would require specification of some functional form. In such cases stronger assumptions are required and the similarity of the reduced form before and after the program will depend on the nature of the functional form assumed. Identifying the *ex ante* treatment effect also requires that any unobserved heterogeneity (ν) remains the same before and after treatment ie. (ν) is independent of consumption and school costs. However, the use of consumption in the reduced form means there is the problem of potential endogeneity. This would particularly be the case if decisions about schooling depend on the opportunity costs of enrolment or if parents are heterogenous in preferences to invest in children's education. These preferences are likely to depend a great deal on family structure and background etc. To make this assumption on unobserved heterogeneity plausible, empirically the matching functions include a set of family characteristics - X_h and price variables.

$$f(\nu|F, \theta, X_h) = f(\nu|\tilde{F}, \tilde{\theta}, X_h)$$

2.4 Empirical Specification

The above approach generates a set of variables that naturally extend to an empirical application of the model. This relies on direct variations in the policy variables. In this case variation in the costs of schooling and full income can be exploited and a matching estimator applied to identify predicted program impacts.

Typical evaluation exercises use information on treated outcomes (S_1) and estimate the counterfactual of untreated outcomes (S_0). In contrast, in the *ex ante* approach treated outcomes are unobserved and is the counterfactual that needs to be estimated. From the model, as indicated by Todd & Wolpin (Forthcoming), the unobserved S_1 can be represented in terms of the observed untreated outcomes conditional on an equivalent set of exogenous variables. This idea can be represented as:

$$S_{1i} = E[S_{0j}|F_i = \widetilde{F}_j, \theta_i = \widetilde{\theta}_j, n_i = n_j, g_i = g_j, X_{hi} = X_{hj}] + \epsilon \quad (2.8)$$

Todd and Wolpin propose a matching estimator of the average treatment effect for those eligible for the program (intent-to-treat (ITT)) as:

$$\alpha = \frac{1}{k} \sum_{j=1}^k \sum_{i \in S_m} E(S_i | F_i = \widetilde{F}_j, \theta_i = \widetilde{\theta}_j, n_i = n_j, g_i = g_j, X_{hi} = X_{hj}) - S_j(F_j, \theta_j, n_j, g_j, X_{hj}) \quad (2.9)$$

2.4.1 Estimating School Costs

Implementing the above matching estimator requires estimation of the unobserved treated outcomes as a function of consumption, school costs and a set of household and child characteristics. School costs (δ_i) are determined by the enrolment status of the child and hence are observed in the data for only those children who are currently enrolled in school and zero costs observed for those not enrolled. The problem of predicting school costs for the entire

sample of children requires using a two-step process decomposing the participation decision and the determinants of the cost of schooling. A *two-part model (2PM)*² is applied where in the first part, the enrolment decision, is modelled using a probit and the second part predicts the cost of schooling as a linear function of the determinants of school costs (Mullahy 1998). The most common specification of the second part is a log transformation of the outcome variable. A problem with using a retransformed OLS in this case is that zero school costs are also observed in the sample of those children currently attending school. A log transformation would drop these observations from the estimation sample. A further problem arises with retransformation of the outcome variable to the original scale in the presence of heteroskedasticity. Manning (1998) shows that heteroskedasticity leads to biased estimates of the outcome variable and correction requires determining whether the heteroskedasticity is across different groups or caused by a particular subset of the covariates. To overcome these issues the second part of the 2PM is estimated using the *extended estimating equations model (EEE)* proposed by Basu & Rathouz (2005). The EEE approach is an extension of a standard *generalized linear model (GLM)* incorporating flexible link and variance functions. Specifically, the EEE combines a Box-Cox transformation for the link function and includes a class of link functions represented by an estimated parameter λ :

$$\frac{\delta^\lambda - 1}{\lambda}$$

It also allows for heteroskedasticity and uses a general power function for the variance defined by two-parameters θ_1 and θ_2 :

$$\theta_1 \delta^{\theta_2}$$

The model is estimated separately for boys and girls.

²A bivariate selection model was initially estimated and the non-linearity of the inverse mills ratio showed no evidence of selection bias. Also, school costs are not normally distributed and a log transformation would drop the observations that indicated zero costs. A further difficulty arises in finding a suitable exclusion restriction that affects school enrolment but not school costs. To overcome these problems a two-part model is used.

2.4.2 Estimating Counterfactual Outcomes

The unobserved outcomes $E(S_i|F_i = \widetilde{F}_j, \theta_i = \widetilde{\theta}_j, n_i = n_j, g_i = g_j, X_{hi} = X_{hj})$ can be estimated using a binary response model to estimate the conditional probability $P(S = 1|X = x) = G(x\beta)$. If the distribution function G is known a priori then a parametric specification such as a logit or probit can be used. Misspecification of G would however result in inconsistent estimates of β and inaccurate predictions of the unobserved outcomes. To increase the flexibility and avoid misspecification problems the unobserved outcomes are estimated by regressing current enrolment status on income, estimated school costs, and a set of family and child characteristics to capture unobserved heterogeneity using a semiparametric *single-index model*. The single-index model defines the conditional mean function as:

$$E(S|x) = G(x\beta) \tag{2.10}$$

where β is an unknown vector and G is an unknown function and $x\beta$ represents an index. The above index specification could be made entirely flexible using a fully nonparametric approach to model outcomes eliminating the risk of any misspecification. Such an approach is however constrained in this case by the dimensionality of the covariate vector (x). Non-parametric approaches suffer from the *curse of dimensionality* where convergence rates are inversely related to the number of continuous covariates and tend to be less precise as the dimension increases. The single-index $x\beta$ reduces the dimensionality problem by aggregating across x and has the same convergence rate as a single dimensional quantity represented by $x\beta$. The single-index model also has advantages for predictions as the region of support extends beyond the observed x to points not in the support of x but in the support of $x\beta$ (Horowitz 1998). However, unlike the nonparametric approach it builds in a parametric assumption of the linearity of the index.

The single-index model involves the joint estimation of the two unknown elements β and G . Estimation of both elements require several identification restrictions. Similar to all linear models, identification of β requires G to be a non-constant function along with the absence

of multicollinearity amongst the covariates. In addition, to uniquely identify the function $G(x\beta)$ single-index models involve *location normalization* and *scale normalization* restrictions. Location normalization is achieved by requiring the covariate vector to include no intercept term while scale normalization involves restricting the β coefficient of one continuous variable to equal one. Identification in single-index models is achieved because the conditional mean function can remain constant with changes in x as long as the index $x\beta$ remains constant. However, with continuous covariates a constant index (ie. $x\beta = k$) for a given set of covariates has probability zero. To overcome this a further identification restriction is required where G is a differentiable function so that $G(x\beta)$ is close to $G(k)$ when $x\beta$ is close to k (Horowitz 1998). A final set of restrictions are required when X contains both discrete and continuous variables. The first of these requires that the discrete elements of the covariate vector do not divide the support of $x\beta$ into disjoint subsets. The final restriction is referred to as the 'non-periodicity condition' for the function G requiring it to be strictly increasing.

The single-index model defined in (2.10) was adapted to binary outcomes by Klein & Spady (1993). In the case of binary outcomes such as enrolment (where $S = 0, 1$) the index function is defined as:

$$E(S|x) = P(S = 1|x) = G(x\beta)$$

In a parametric setting with known G , β could be estimated efficiently using a maximum likelihood estimator(MLE) where the log-likelihood is:

$$\ln L(\beta, G) = n^{-1} \sum_{i=1}^n [S_i \ln G(x_i\beta) + (1 - S_i) \ln(1 - G(x_i\beta))] \quad (2.11)$$

In the semiparametric case following Ichimura (1993), Klein and Spady propose to estimate β by maximising the (quasi) log-likelihood function(2.11) replacing the unknown function G with a semiparametric likelihood estimate $G_n(x_i\beta)$. 'The index restriction permits multiplicative heteroskedasticity of a general but known form and heteroskedasticity of an unknown form if it depends only on the index' Klein & Spady (1993). G_n is estimated using a

leave-one-out nonparametric estimator of the density of $x\hat{\beta}$ conditional on S , where for any z

$$G_n(x_i\beta) = \frac{P_n g_n(z|S=1)}{P_n g_n(z|S=1) + (1 - P_n) g_n(z|S=0)} \quad (2.12)$$

where g_n is the kernel estimate of the conditional density of $x\beta$ ($g(\cdot|S)$) and g_n is defined as:

$$g_n(z|S=1) = \frac{\sum_{i=1}^n S_i K(z - x_i\hat{\beta})/h_n}{nP_n h_n} \quad (2.13)$$

$$g_n(z|S=0) = \frac{\sum_{i=1}^n (1 - S_i) K(z - x_i\hat{\beta})/h_n}{n(1 - P_n) h_n} \quad (2.14)$$

where P_n is the empirical probability $P_n = \sum_{i=1}^n S_i$, the proportion of children currently enrolled in school, K is a kernel function and h_n is the bandwidth.

Klein and Spady show that the estimator is asymptotically efficient and achieves the semi-parametric efficiency bounds of Chamberlain (1986) and Cosslett (1987). The resulting vector of parameter estimates ($\hat{\beta}$) is shown to have the following properties:

$$n^{1/2}(\hat{\beta} - \beta) \rightarrow_d N(0, \Omega)$$

$$\Omega = E \left\{ \left[\frac{\partial G(X_i\beta)}{\partial \beta} \right] \left[\frac{\partial G(X_i\beta)}{\partial \beta} \right]^T \left[\frac{1}{G(X_i\beta)(1 - G(X_i\beta))} \right] \right\}^{-1}$$

The unobserved treated outcomes $E(S_i|F_i = \tilde{F}_j, \theta_i = \tilde{\theta}_j, n_i = n_j, g_i = g_j, X_{hi} = X_{hj})$ are estimated using the Klein and Spady estimator by regressing school enrolment status of

the observed “control group” children on observed consumption, school costs and a vector of child and household characteristics. Scale normalization is achieved by setting the coefficient for number of children under 5 years in the household equal to 1. The estimated model is then used to first predict enrolment outcomes for the observed control group observations and then extrapolate the predictions under treatment by evaluating the function at $(\tilde{F}, \tilde{\theta})$.³

Both within sample predictions and extrapolation can only be carried out in regions of common support. In the original formulation of the model, as required by the QMLE asymptotic theory, Klein and Spady introduce trimming procedures on the likelihood function (14) to ensure that G is bounded away from 0 and 1. But their simulations show that trimming has little impact in empirical applications. Following their findings and other applications of this model (Horowitz 1993, Gerfin 1996, Fernández & Rodríguez-Poo 1997) the likelihood function is not trimmed before predicting outcomes for the observed data. Extrapolation in nonparametric models is only valid at points with positive data density; hence while not trimming the likelihood function, trimming is carried out to define the region of common support i.e to identify regions of positive data density in the extrapolated values. The region of support S_m is defined as $S_m = \{x\beta \in R^2 \text{ such that } f(x\beta) \geq c\}$ where $f(x\beta)$ is the non-parametric density of the linear index⁴. Heckman, Ichimura & Todd (1997) propose that the density should be strictly positive as defined by S_p and should exceed a minimum cut-off to avoid points with very low density. Thus the extrapolation is valid for only those points of evaluation where

$$f(x\hat{\beta}) > c \tag{2.15}$$

Heckman, Ichimura & Todd (1997) recommend setting the cut-off at a percent quantile of the estimated densities. Here c is set at the 2% quantile. Only those observations that meet the above criterion are kept in the extrapolation sample.

³The statistical package np (Hayfield & Racine 2008) available for the software R was used. The model was run separately for boys and girls. The scalar bandwidth for the index $x\beta$ for boys is 0.083 and for girls is 0.065.

⁴The densities are estimated using the method of Li & Racine (2003) who use ‘generalized product kernels’ for mixed data. The bandwidths were set using the maximum likelihood cross validation

2.5 Data and Variables

This paper uses data collected for the *ex post* randomized evaluation of RPS. Two datasets (IFPRI 2005) from the *ex post* evaluation are applicable, the first is the census survey conducted in May/June 2000 covering all eligible households in the two regions selected for the program and the second, the baseline survey in August/September 2000, conducted for the randomized experiment prior to introduction of the subsidies. The data in the baseline survey includes detailed information on school enrolment, detailed direct and indirect costs (including fees, transport, books, uniforms, etc.) on schooling for those enrolled; health care utilization including consultations, type of provider, use of medication and hospitalization, direct and indirect costs of medical care and waiting times. However, the information on economic activity is sparse with only information on employment status, nature of employment, category of employment and hours worked. No information was collected on wages or income. All the above information was collected for all individuals of age 6 and over. Lack of information on income is substituted by information consumption.

2.5.1 Variables

The census data provides information on the highest grade and level of education completed by all individuals aged 6 and over. The education of the household head can be mapped from this to the baseline survey. The census survey is also useful in trimming the sample to the program eligible children between the ages of 7-13 who have not completed grade 4. It also provides information on the distance to the nearest primary and secondary school.

School costs are only observed in the data for those children currently enrolled in school and must be estimated for all children in the sample. As mentioned above, both direct and indirect costs are observed and are aggregated into a single measure of costs. Human capital theory bases the family's choice of schooling on costs - both direct and indirect (opportunity costs), income and future returns to education (Becker 1975). School costs are estimated using variables that capture these factors and include child characteristics - age of the child and

gender. The lack of wage data poses a problem in estimating time costs of schooling. To overcome this, distance to school is used as a measure of opportunity cost of travel time. Family characteristic variables such as household expenditure, age, gender and years of schooling completed for the household head, number of children of school going age and number of adults in the family are included. Additionally, number of children under 5 years is used as a measure of demand for child labour as often older children are expected to care for younger siblings.

The estimation of the unobserved schooling outcomes under treatment $E(S_i|F_i = \tilde{F}_j, \theta_i = \tilde{\theta}_j, n_i = n_j, g_i = g_j, X_{hi} = X_{hj})$ is driven by the variables in the reduced form equations derived by the economic model ie. the variables determining the schooling decision are derived from $S^{**} = \Phi(\tilde{F}, \tilde{\theta}, n, g; \nu)$ and include consumption, a quadratic specification of age, whether the community is a coffee growing community, estimated school costs, years of education and gender of the household head, household composition and time and opportunity cost related variables - distance to nearest primary and secondary schools and public transport.

The baseline data covers 9747 individuals (both treatment and control) for 1581 households. This evaluation focuses on outcomes of children eligible for the schooling component of the program. Such households receive both the food transfer and the education transfer components of the program. The sample size for the purpose of this evaluation consists of 1786 children. Over half of this sample consists of families with more than one child eligible for the program.

2.6 Results

2.6.1 Estimating School Costs

Table 2.1 shows the results from estimating the two part model for boys and girls. The probit participation model for both boys (1) and girls (3) show a similar pattern, with enrolment

Table 2.1: Estimating School Costs

VARIABLES	(1) Probit-Boys Enrollment	(2) EEE-Boys School Costs	(3) Probit-Girls Enrollment	(4) EEE-Girls School Costs
age8	0.116 (0.151)	0.117 (0.139)	0.251 (0.160)	0.168 (0.103)
age9	0.265 (0.161)	0.180 (0.147)	0.544** (0.168)	0.385*** (0.103)
age10	0.174 (0.164)	0.0439 (0.124)	0.285 (0.172)	0.345* (0.172)
age11	-0.00216 (0.164)	0.0458 (0.133)	0.202 (0.179)	0.470*** (0.116)
age12	-0.00164 (0.172)	0.189 (0.137)	0.147 (0.188)	0.523** (0.197)
age13	-0.554*** (0.168)	0.0256 (0.138)	-0.159 (0.198)	0.205 (0.134)
HH Consm (adjusted)	0.0000116** (0.00000447)	0.0000405*** (0.00000345)	0.0000122* (0.00000593)	0.0000264*** (0.00000359)
School dist	-0.00703*** (0.00178)	0.00346** (0.00118)	-0.00806*** (0.00173)	0.00230 (0.00155)
No. of adults	-0.0336 (0.0328)		0.0536 (0.0413)	
Children under5	-0.172** (0.0550)	-0.224*** (0.0398)	-0.228*** (0.0621)	-0.0780 (0.0430)
Children 7-13	0.0814 (0.0463)	-0.258*** (0.0443)	-0.0462 (0.0557)	-0.240*** (0.0391)
HHH gender	0.335 (0.183)		-0.0755 (0.198)	
HHH age	0.00714 (0.00529)		0.00979 (0.00562)	
HHH yrs of ed	0.106** (0.0381)		0.180*** (0.0436)	
HHH works	-0.0949 (0.173)		0.186 (0.194)	
Constant	-0.0227 (0.325)	-0.377* (0.155)	-0.0752 (0.375)	-0.288* (0.131)
λ		0.289* (0.143)		0.663** (0.204)
θ_1		1.242*** (0.0887)		1.564*** (0.157)
θ_2		1.597*** (0.106)		1.737*** (0.111)
Observations	945	687	845	631

Robust standard errors in parentheses, clustered at the household level

*** p<0.01, ** p<0.05, * p<0.1

being most likely between the ages of 8 and 10 as compared to children aged 7 (reference category) and declining with older children. Boys drop out earlier (above age 10) while girls aged 13 are less likely to enrol when compared to the reference group. This pattern follows most developing countries where many children enrol and stay in school only for a few years, dropping out between the ages of 11-13 to find employment. Consumption net of school costs (used as a proxy for income) and education of the household head are significant and have a positive impact on enrolment. As mentioned earlier the probit model includes the number of children under 5 years as a proxy for child labour. The estimates show similar negative magnitudes for boys and girls indicating having younger children in the household decreases the likelihood of enrolment. A similar effect of distance to the nearest school is observed, with children being less likely to enrol if schools are further away. Enrolment probabilities differ for boys and girls depending on the gender and the employment status of the head of the household. Girls are less likely to enrol if a male is head of the household, as is the case in 88% of the households in the sample. The direction of the coefficient for employment status is less intuitive as boys seem less likely to enrol if the household head is employed. This result is probably due to the nature of employment, with about 85% of the sample being involved in farm activities. The last two variables though not significant in the model do indicate the presence of a gender gap from additional opportunity costs for boys and cultural differences that contribute to the differences in schooling.

Columns (2) and (4) of Table 2.1 provide results from the second part of the two part model using the *extended estimating equations model (EEE)* (Basu & Rathouz 2005) for school costs.⁵ Boys in the reference category (age 7) face the highest school costs. At other ages there is no significant impact on school costs. For girls however, school costs increase with age. Families with greater consumption tend to spend more on education, although more on the boys than the girls. In both cases children of the same age and children under five is significant (except for girls -children under5) and negative. This is intuitive in the sense that sharing of resources reduces the costs per child as the number of school age children increases.

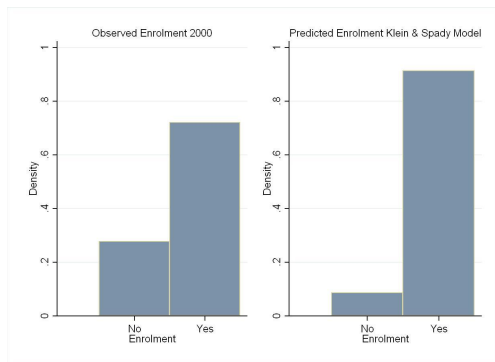
⁵An alternative approach to the EEE model would be to use a *generalized linear model* with a specified link function and distribution. However, failure to specify the correct link function results in misspecification of the model. To avoid such misspecifications, the EEE approach was used since it does not require an *a priori* assumption of a link function or distribution. This approach 'helps to identify an appropriate link function and to suggest an underlying distribution for a specific application but also serves as a robust estimator when no specific distribution for the outcome measure can be identified' Basu & Rathouz (2005).

In Column (2) for the boys sample the link parameter is estimated to be $\lambda = 0.289$ (95% C.I: 0.01, 0.57). The variance function represented by $\theta_1 = 1.2$ (95% C.I:1.07 ,1.42) and $\theta_2 = 1.6$ (95% C.I:1.39 , 1.80) is close to a gamma distribution. Column (4) provides the estimates for the sample of girls. In this case with $\lambda = 0.66$ (95% C.I: 0.26, 1.06), the link function is close to a square root link. The values $\theta_1 = 1.5$ (95% C.I:1.26 ,1.87) and $\theta_2 = 1.74$ (95% C.I:1.51, 1.95) again suggest a gamma distribution.

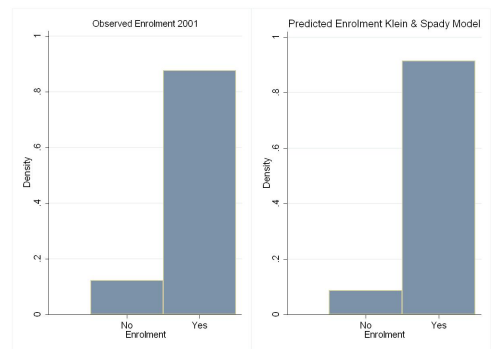
2.6.2 Predicting Impacts

The empirical specification of the Klein and Spady model described in section 2.4 is used to predict unobserved school enrolment ($S^{**} = \Phi(\tilde{F}, \tilde{\theta}, n, g; X_h)$) under the treatment, accounting for the age of the child and a quadratic specification of age, number of children under 5, number of children between 7 and 14 years, number of adults, years of education and gender of the head of the household, the household lives in a coffee growing community, consumption, school costs, distance to the nearest primary and secondary school and public transport. Figures 2.4(a) and 2.4(c) illustrate the observed data from the comparison group (S_j in equation 2.9) along with the Klein and Spady predictions for the extrapolated unobserved outcomes under treatment (S_i in equation 2.9). Figures 2.4(b) and 2.4(d) compare predicted outcomes from the Klein and Spady model with those observed in the 2001 follow-up survey of the experiment. A comparison of Figures 2.4(b) and 2.4(d) shows that the extrapolated outcomes are quite close to the observed follow-up data for both boys and girls.

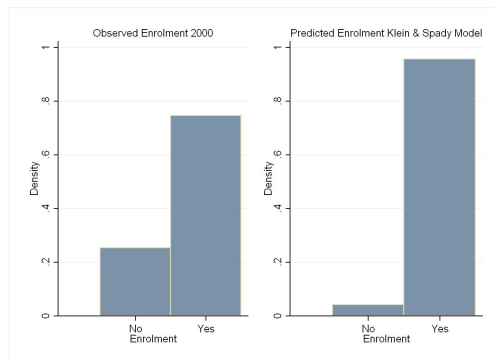
The estimator in equation (2.9) matches baseline program eligible children with characteristics $(\tilde{F}, \tilde{\theta}, X_h)$ with other baseline program eligible children with characteristics (F, θ, X_h) . The estimated treatment effect is only valid for those families within the region of common support defined by equation (2.15). Figures 2.5(a), 2.5(b), 2.5(c) and 2.5(d) compare the distributions of the variables included in the matching before and after trimming is implemented in the Klein and Spady estimator. They show that as required trimming eliminates observations where the density is very low, this translates to the ends of the right-tail of the distributions i.e families with very high consumption or school costs for whom matches are unlikely to be available are dropped from the estimation of treatment effects.



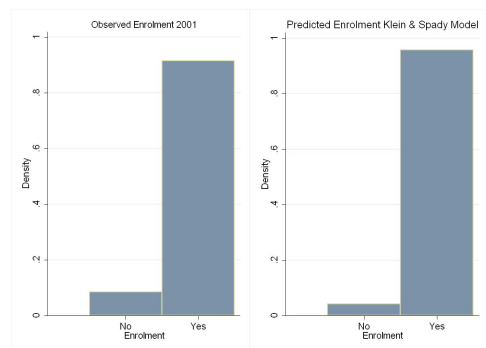
(a) Observed Baseline Outcomes with Predicted Outcomes for Sample of Boys



(b) Observed Follow up Outcomes with Predicted Outcomes for Sample of Boys



(c) Observed Outcomes and Predicted Outcomes for Sample of Girls



(d) Observed Follow up Outcomes with Predicted Outcomes for Sample of Boys

Figure 2.4: Comparing Observed and Predicted Outcomes

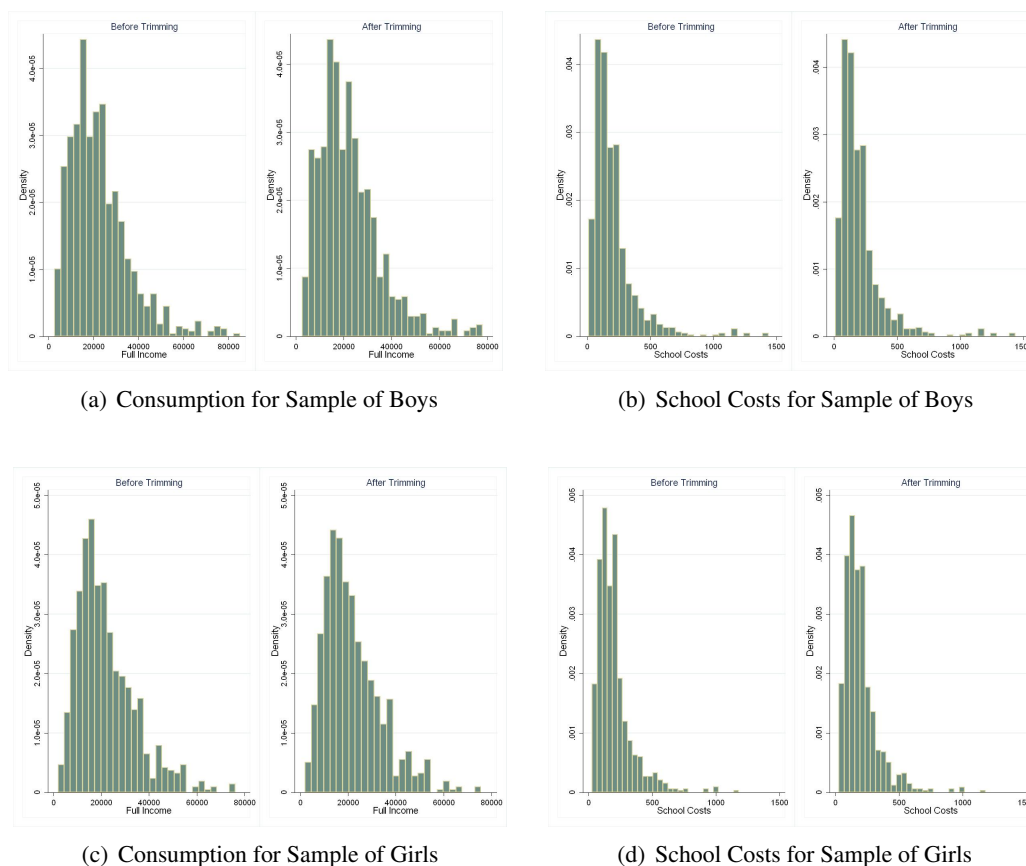


Figure 2.5: Trimming Klein and Spady estimations

The predicted impacts are listed in column (1) of Table 2.2 along with corresponding results from the *ex post* evaluation of RPS (column 3) ⁶. This paper evaluates the impact of RPS using both treatment and control group data at the baseline as a single cross-section rather than just control group data as in the case of Todd & Wolpin (Forthcoming). This is in some ways similar to choosing a sample of program eligible households from any available cross-section (such as the Living Standards Measurement Survey (LSMS)). A comparison of the *ex ante* and *ex post* outcomes show that the *ex ante* approach predicts very closely the overall program impact for both boys and girls and is statistically significant, with one year of conditional cash transfers having a positive impact on enrolment of both boys and girls. The estimated impact for boys is 0.19 and accurately predicts the results of the ex-post evaluation. The one-year cash transfer increased enrolment of girls by 21 percentage points as compared to 20 percentage points from the experimental evaluation. In comparing the enrolment between boys and girls, girls continue to have higher enrolment rates even after 1

⁶The ex-post evaluation results from the published report of the evaluation of RPS provide only the overall impact. The other values were calculated for this evaluation.

Table 2.2: Predicted Impact

VARIABLES	(1) Predicted Impact	(2) Sample sizes [@]	(3) Experimental impact
Boys 7 -13	0.19*** (0.0222)	859 / 876	0.19***
Girls 7 -13	0.21*** (0.0219)	754 / 767	0.20***
Boys & Girls 7-9	0.17*** (0.0255)	829 / 844	0.23***
Boys & Girls 10-13	0.15*** (0.0316)	786 / 799	0.15***

- (1) The predicted and experimental results are the combined effects of all three transfer components of the program Bono escolar (school attendance), Mochila escolar (school supplies) and Bono alimentario (food security).
(2) @ treatment observations after trimming, total number of observations.
(3) Bootstrapped standard errors clustered at the comarca level (500 reps).
(4) *** p<0.01, ** p<0.05, * p<0.1

year of the program. At the baseline 75% of program eligible girls were already enrolled in school as compared to 72% of boys. The difference in opportunity costs could be a factor in explaining this difference as girls could be more likely to enrol possibly due to lower opportunity costs as compared to boys.

To examine further the predictions, the impacts are analysed by subgroups of age. The same model specification used for the boys and girls is used to estimate *ex ante* impacts for two sub-groups- children below 10 and those 10 years old and above. The *ex ante* evaluation estimates an impact for children 10 years old and above as .15 which accurately predicts the experimental results. In the case of children below 10 years, the experimental evaluation shows a very large 23 percentage point rise in enrolment. The *ex ante* estimates for the same age group are also large and positive but are lower in magnitude at 17 percentage points when compared to the experimental estimates. The method relies on accurately capturing all direct and indirect costs of schooling to reflect the price effect of receiving the cash transfer for the school component of the program. The survey provided no information on time use and earnings of children. The underprediction could reflect children being involved in work at home outside of their school hours or in non-paid employment in the household. The lack of more detailed information could be leading to the underprediction. In general the RPS program shows large and positive impacts across boys and girls and for different age groups.

However, the impacts on children of younger ages is greater than for older children. The reduced form behavioural model approach applied here performs well in predicting one-year *ex ante* impacts of the RPS program.

2.7 Conclusion

This paper presents an *ex ante* evaluation of Nicaragua's CCT program Red de Protección Social. It applies the methods proposed by Todd & Wolpin (Forthcoming) on using reduced form estimation of behavioural models to carry out *ex ante* evaluations of social programs. The key requirement in this approach is that the preferences remain the same before and after the program so that the impact of the program is captured by a change in the magnitude of the exogenous variables resulting from an introduction of the program. This paper extends their approach to model education utilization outcomes and presents results from the schooling component of the program.

The model considers the influence of both direct and opportunity costs of schooling. Variation in the policy variable (school costs) and full income of the household is exploited to estimate program impacts on school enrolment. Empirically the model is implemented using a semi-parametric single index framework that allows for an increase in the dimensionality of the covariate vector. The outcome, school enrolment, is binary and the semi-parametric estimator proposed by Klein and Spady is used to predict the unobserved outcomes under treatment. The data set combines baseline data from the RPS experiment along with some information from the census survey . The baseline data is used as a single cross-section combining both control and treatment groups. Comparing the predicted estimates with the experimental outcomes shows that the predictions all have the same direction as the experimental impact. The predictions for overall impact of the program for boys, girls and the age group - 10 years and above are very close in magnitude to the experimental impact. The prediction for younger children however is lower than the experimental impact but still shows a large positive effect of a one year cash transfer. The empirical approach used relies on selection on observables and performs well when the observables are fully captured. In general, in keeping with the findings from the experiment, the *ex ante* evaluation finds a significant and large overall impact of RPS on the target population.

Chapter 3

Conditional Cash Transfers to Improve Education and Health: An Ex ante Evaluation of Red de Protección Social, Nicaragua

3.1 Introduction

Millions of children under 5 years face risks to their development from poverty, poor nutrition and limited access to healthcare. Grantham-McGregor et al. (2007) estimate that the above risks prevent more than 200 million children under 5 years from reaching their potential in cognitive development. Poor nutrition, untreated infections and lack of growth monitoring can result in growth retardation and stunting. Evidence from several developing countries shows that these children are in turn less likely to enrol in school and perform as well as children not facing these risks (Moock & Leslie 1986, Brooker et al. 1999, Glewwe et al. 2001). This lack of investment in human capital results in intergenerational transmission of poverty and poor health.

Early interventions that promote child development ameliorate the above risks and are critical to improving health outcomes and breaking the circle of poverty (Engle et al. 2007). To address the problem conditional cash transfer(CCT) programs have emerged as popular social protection schemes and are now implemented by many governments in Latin America and several parts of Africa. While these programs have largely focused on improving access to education for children, some programs have also implemented health components with a

view to reducing current poverty while encouraging investment in human capital and health. Typically the health components of these programs look to improve access to preventive interventions and monitor child growth.

Evaluating the impact of these programs is critical to understanding their effectiveness in different settings. A few of the CCTs in Latin America were implemented as pilot randomized experiments. While these experiments provide the best estimate of the *ex post* treatment effect, they are limited in their application to exploring alternative policy scenarios or estimating the impact of expanding a pilot program. Such analyses require *ex ante* evaluations using behavioural models. Two approaches are available to carry out *ex ante* evaluations. The first approach relies on structural estimation of a dynamic behavioural model. Todd & Wolpin (2006) and Attanasio et al. (2005) apply this approach to estimate a discrete choice dynamic programming model of education and fertility outcomes for the Progresa program in Mexico. More recently, building on the work by Ichimura & Taber (2000), Todd & Wolpin (Forthcoming) propose a second approach based on estimating reduced form equations using minimal assumptions of functional form and applying non-parametric estimation methods to compare predicted school enrolment outcomes with those of the experiment for the same program. The present paper tests their reduced form approach and explores the feasibility of this approach in forecasting health related outcomes. For the health related outcomes it relies on a health production framework and considers the impact of the cash transfer on demand for immunizations and the take up of child health checks. The paper also carries out simulations of alternate policy scenarios.

The data used in this paper is from the Red de Protección Social CCT program in Nicaragua. The program was implemented in 2000 as a randomized social experiment and has two equally large cash transfer components, one for school enrolment and the second for food security and nutrition. The behavioural model reduced form (BMRF) approach used in this paper differs from Todd & Wolpin (2006) and Attanasio et al. (2005) in that it does not estimate the structural parameters of the model. Instead it relies on variation in the observed policy related variables, and models the reduced form demand equations for the outcome variables from an underlying static model. Identification in this approach requires the program to channel its influence solely through the budget constraint and the reduced form functions

before and after the program remain the same. Empirically, the model is implemented using a semi-parametric single index model which relaxes the limitation on the number of exogenous variables possible in a fully non-parametric approach. The semi-parametric approach also allows for greater flexibility in capturing confounders.

3.2 Background

3.2.1 CCTs and health services

CCT programs are social protection programs aimed at breaking the intergenerational transmission of poverty and poor health. These programs offer cash transfers that can be spent without restrictions, however receiving the cash transfer is conditioned on altering certain behaviours - such as enrolment of children in school and investing in preventive health care for children. The goals are to simultaneously reduce poverty while encouraging investment in human capital. The cash transfers are a form of subsidy that reduce the financial barriers to accessing either education or health services. The level of services accessed by a household are driven by both demand and supply side factors, where on the demand side the level and frequency with which an individual or household accesses care depends on their level of informed decision making i.e - the ability with which they are able to identify being ill or the need for preventive care and the capacity to access and utilize care (Ensor & Cooper 2004). CCT programs target these demand side barriers. Three assumptions underlie the health components of CCT programs (Glassman et al. 2007); the first, that poor households under-utilize health services, the second, that these families do not possess sufficient health education and knowledge about the benefits of accessing preventive care and third, in order for the transfer to have an impact receipt of the transfers must be conditioned on some health related behaviour altering requirement.

3.2.2 Health care utilization and vaccinations

Health services in a CCT program targeted at children are usually provided as a package of services with the cash transfer serving as an incentive to access the package. CCTs recognize

the production of child health has several inputs, including maternal schooling and endowments (abilities of the mother, knowledge of nutrition and habits) and access to preventive care (Behrman & Wolfe 1987). Embedding different outcomes in a package of services could increase demand for individual outcomes more than if they are provided as separate services as the marginal impact of one input may be higher when combined with the other relevant inputs (Strauss & Thomas 2008).

One major outcome, that in most CCT programs is designed as an explicit conditionality, is health checks for children below 3 years. The use of health checks for children is an important input in overall health and development. Accessing these services is however closely associated with the level of education of the parents and the time and financial costs of accessing care. Thus, by making health checks for children a conditionality for the cash transfer, these programs look to minimize the impact of a family's socioeconomic status on child health seeking behaviour. Evidence from *ex post* evaluations of CCT programs from three Latin American countries - Honduras, Nicaragua and Colombia (see Glassman et al. (2007) for a summary) have shown significant increases in the number of children being taken to health clinics.

Vaccinations against preventable diseases such as measles, tetanus and whooping cough are the most cost-effective of preventive health interventions (Miller & Hinman 2004, Hadler 2004) and are provided within the package of services provided by CCTs. However, unlike health checks, they are rarely made an explicit conditionality and are typically provided under the general set of services provided at health checks. The *ex post* evidence is limited due to the lack of complete data for the different vaccines that are given to achieve "full immunization". One exception is the RPS program in Nicaragua. The randomized evaluation (Barham & Maluccio 2009) shows a 19 percentage point increase in full coverage as a result of the program. Evidence from the Progres program in Mexico and the CCT program in Honduras is limited to certain vaccines, but both find small but positive impacts.

3.2.3 Red de Protección Social

Red de Protección Social (RPS) was implemented in 2000 as a pilot randomized evaluation in 42 localities of six rural municipalities of Nicaragua. The pilot evaluation was maintained

as an experiment for two years following which the initial control group was integrated with the treatment group and provided the cash transfers. The program was designed to target education of children in rural households and had two transfer components. The first was a school transfer component given conditional on all children in the household between the ages of 7-13 who have not completed grade 4 of primary school enrolling and maintaining 85% attendance. Independent of the school component the food security, health and nutrition transfer was provided directly to mother's of beneficiary households conditional on (1) bringing her children to scheduled preventive health checks, (2) attending bimonthly health education workshops and (3) adequate weight gain for children. Figure 3.1 presents a summary of the eligibility criteria and requirements for RPS.

The health services provided during the compulsory checks included growth monitoring, vaccinations and nutrition supplements for anaemia and anti-parasite medicines. Beneficiaries were required to use RPS trained and enlisted providers. Providers enlisted by RPS were paid to travel to program regions by the program operators to provide vaccinations and the additional health services in existing facilities or other community facilities. The required infrastructure and stock of medicines and vaccinations were monitored and provided at these facilities by the program operators. However, there was a delay in expanding the government health services to meet the demand from the program and hence the health related conditionalities were not enforced during the first 8 months of the transfers. This however means that the experimental results are a combination of different aspects which cannot be isolated ie. the result combines the effect of the cash transfer and the conditionalities (for the last four months). In which case the ex ante results from this paper represent the effect of only the cash transfer component. This point is discussed further in the results section.

The amounts of the transfer for each household included, the Córdoba 2000 equivalent of US\$224 for food security and US\$112 for the education component per year. In addition families also received a per child school supplies transfer of US\$21. Figure 3.2 shows the summary of the transfer amounts. Maluccio & Flores (2005) estimate that the food transfer was approximately 13% of annual household expenditure and if families had only one school component eligible child, then they received an additional 8% of annual household expenditure.

PROGRAM REQUIREMENT	HOUSEHOLD TYPE		
	Households with no targeted children (A)	Households with children aged 0-5 (B)	Households with children aged 7-13 who have not completed 4 th grade (C)
Attend bimonthly health education workshops	✓	✓	✓
Bring children to prescheduled healthcare appointments Monthly (0-2 years) Bimonthly (2-5 years)		✓	✓
Adequate weight gain for children under 5 ^a		✓	✓
Enrollment in grades 1 to 4 of all targeted children in the household			✓
Regular attendance (85 percent, i.e., no more than 5 absences every two months without valid excuse) of all targeted children in the household			✓
Promotion at end of school year ^b			✓
Deliver teacher transfer to teacher			✓
Up-to-date vaccination for all children under 5 years ^b		✓	✓

a. The adequate weight gain requirement was discontinued in Phase 11, starting in 2003

b. Condition was not enforced.

Figure 3.1: RPS Eligibility and Requirements. Source: Maluccio and Flores 2005

Nicaraguan RPS eligibility and benefits in Phase 1		
	Program components	
	Food security, health and nutrition	Education
Eligibility		
Geographic targeting	All households	All households with children aged 7-13 who have not completed fourth grade of primary school
Demand-side benefits		
Monetary transfers	<i>Bono alimentario</i> (food security transfer) C\$2,880 per household per year (US\$224)	<i>Bono escolar</i> (school attendance transfer) C\$1,440 per household per year (US\$112) <i>Mochila escolar</i> (school supplies transfer) C\$275 per child beginning of school year (US\$21)
Supply-side benefits		
Services provided and monetary transfers	Health education workshops every 2 months Child growth and monitoring Monthly: Newborn to 2-year-olds Every 2 months: 2- to 5-year-olds Provision of antiparasite medicine, vitamins, and iron supplements Vaccinations (newborn to 5-year-olds)	<i>Bono a la oferta</i> (teacher transfer) C\$80 per child per year given to teacher/school (US\$6)

Figure 3.2: RPS Transfers. Source: Maluccio and Flores 2005

3.3 Theoretical Framework

This paper focuses on empirically estimating and predicting health related outcomes under the RPS cash transfer program. It relies on a static behavioural model of infant health production (below 3 years) and school enrolment of school aged children (7-13 years).

Health of an infant is assumed to be produced by a production function specified as a function of health related consumption inputs, medical care inputs and a vector of household characteristics.

The health production function is:

$$H_i = h(\mu, E^m, E^m, M; X_h) \quad (3.1)$$

where E^m is a binary indicator of attendance at the conditional educational workshops by

the mother, μ represents additional consumption of nutritional food or other health inputs which is assumed to depend on the transfer received conditional on the mother attending the workshop. This term is set equal to zero in the pre-program situation as E^m equals zero before the program. M represents nutritional and medical care inputs and X_h represents a vector of household characteristics.

A household with multiple eligible children $i = 1 \dots n$, has utility U a function of C representing non-medical consumption, health status of children H , a vector S of binary indicators of school enrolment, with $S_i = 1$ indicating school enrolment, and an indicator of gender g . The household maximisation problem is then:

$$\max_{C, H, S} U(C, H, S; g, \nu) \quad (3.2)$$

The time constraint for a school component eligible child can be written as:

$$T_i = T_{si} \cdot S_i + T_{wi}(1 - S_i) \quad (3.3)$$

where T_{si} is time spent in school and is assumed to be a fixed amount for all enrolled children, T_{wi} is time spent at work. This specification follows Todd and Wolpin and does not allow for leisure in the time constraint as a substitute. However, a different form of time constraint that does not assume substitution is possible, that allows total time T_i to be divided between time spent on work, schooling and leisure. The reduced form in this case would be different to the one derived here.

The money budget constraint of the household can be written as:

$$C + \sum_{i=1}^n \delta_i \cdot S_i + p_m \cdot M - \mu \cdot E^m = Y + \sum_{i=1}^n w_i \cdot T_{wi}(1 - S_i) \quad (3.4)$$

Where $\mu \cdot E^m = 0$ in the pre-program scenario, δ_i is the direct cost of schooling for child i . Primary schooling is free in Nicaragua and most children face no tuition fees, hence δ_i includes all other school related costs faced by families such as transport, uniforms, books

and school meals. p_m is the cost per unit of medical care consumed and Y is household income net of the earnings of the program eligible children. With $\mu.E^m = 0$ in the pre-program scenario the money budget constraint is:

$$C + \sum_{i=1}^n \delta_i \cdot S_i + p_m \cdot M = Y + \sum_{i=1}^n w_i \cdot T_{wi} (1 - S_i) \quad (3.5)$$

The constrained household maximisation problem is:

$$\max_{C, H, S} U(C, H, S; g, \nu) \quad (3.6)$$

which is maximised subject to a full income constraint that combines the time constraint of school going children and the money budget constraint of the household:

$$C + \sum_{i=1}^n [\delta_i + w_i \cdot T_{si}] S_i + p_m \cdot M = Y + w_i \cdot \sum_{i=1}^n T_i = F \quad (3.7)$$

F is full income of the household. The full income constraint expresses the total price of schooling for all eligible children in the family ($\theta = \sum_{i=1}^n [\delta_i + w_i \cdot T_{si}]$) as the costs of schooling plus the shadow wage for the eligible children.

Optimising the utility with respect to the constraints gives the standard reduced form demand functions for the outcomes of interest - schooling $S^* = \Phi(F, \theta, p_m, n; g, \nu)$, health is $H^* = \Omega(F, \theta, p_m, n; g, \nu)$ and preventive care inputs is $M^* = \Psi(F, \theta, p_m, n; g, \nu)$

The RPS program has two cash transfers - the first focuses on changing the price of schooling for eligible children conditional on enrolment and the second is a food transfer meant to boost

consumption, nutrition and access to preventive health care conditional on mother's attending the health workshops (E_m). The initial objective of RPS was to condition the food transfer on a series of other requirements including taking children under 5 years for health checks and maintaining up-to-date immunization. But as explained in the program description these conditionalities were not enforced till almost the second year of the program and hence does not affect the analysis in this paper. The household food transfer (μ) conditional on E_m is modelled as a direct income effect, raising the income level of the household and does not stipulate specific expenditure categories. Under the assumption of full compliance, i.e. all mothers attending the compulsory health education workshops and all children being taken to scheduled health checks, the food security transfer is added to the full income of the household along with the per child transfer for school supplies. By relaying the impact of the program through the budget constraint this approach models the impact of the cash transfer component of the policy. It does not allow for the impact of the different health related conditionalities on health outcomes. At subsistence consumption levels, an increase in income through a transfer is assumed to impact food consumption changing consumption patterns to more nutritious components in the food basket and reducing financial barriers to utilizing preventive care. The focus on health checks and immunization as outcomes without their being implemented as a conditionality is useful in analysing the short-term impact of increased economic status on accessing child health services.

The school transfer is implemented as two components (τ, ρ) to reduce the net price of schooling and substitute for any wages earned by children not enrolled in school due to employment. A decrease in the price of schooling is likely to encourage children to substitute away from labour market participation and increase school enrolment. The first component τ is provided for each eligible child in the family while ρ is a lump sum transfer irrespective of the number of eligible children. Both transfers are conditional on all eligible children enrolling in school.

With the introduction of the subsidies $\mu.E^m$, $\tau.\sum_{i=1}^n S_i.S_p$ and $\rho.S_p$, where $S_p = 1$ if $\sum_{i=1}^n S_i = n$ ie. all eligible children enrol in school and $S_p = 0$ otherwise, the full income for a beneficiary family is:

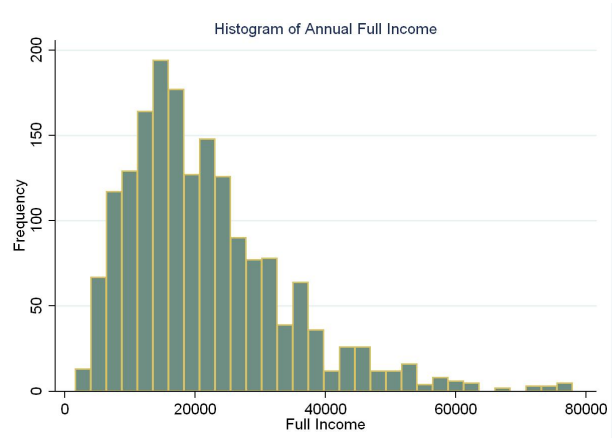
$$C - \rho.S_p - \mu.E^m + \sum_{i=1}^n (\delta_i + w_i.T_{si} - \tau.S_p) S_i + p_m.M = Y + \sum_{i=1}^n w_i.T_i = \tilde{F} \quad (3.8)$$

The new price of schooling under the subsidy program is $\tilde{\theta} = (\sum_{i=1}^n [\delta_i + w_i.T_s - \tau.S_p])$ and the new level of full income is \tilde{F} . The optimal choice under the subsidies is $S^{**} = \Phi(\tilde{F}, \tilde{\theta}, p_m, n; g, \nu)$, health is $H^{**} = \Omega(\tilde{F}, \tilde{\theta}, p_m, n; g, \nu)$ and preventive care inputs is $M^{**} = \Psi(\tilde{F}, \tilde{\theta}, p_m, n; g, \nu)$

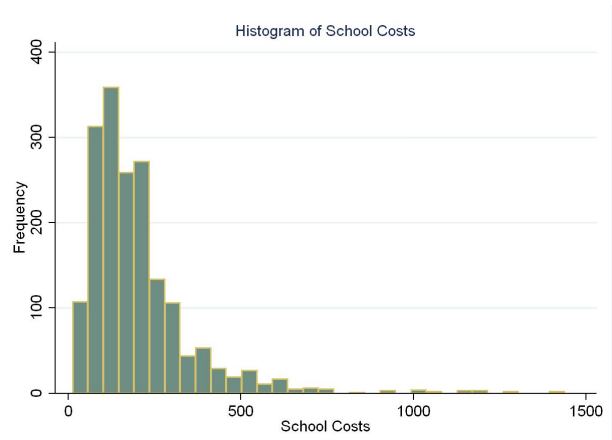
Identifying the ex ante impact of the program using the approach of Todd & Wolpin (Forthcoming) requires the health outcomes reduced form demand function H (and school outcomes function S) to remain the same before and after the program is introduced ie.

$$M^{**} = \Psi(F, \theta, p_m, n; g, \nu) = \Psi(\tilde{F}, \tilde{\theta}, p_m, n; g, \nu) \quad (3.9)$$

The above equation shows that the reduced form functions before and after the program are the same except for the magnitudes of the exogenous policy variables. This assumes that the program has an influence only through the budget constraint and does not directly enter the utility function. Empirically this allows exploitation of two sources of variation in the data to compare untreated individuals with outcomes M^* with other untreated individuals with outcomes M^{**} - the first is school costs and the second is full income of the households at the baseline. As described earlier, primary education is free in Nicaragua and most families face no fees, the cost here includes other expenditure related to schooling which is exogenous in the sense that it is faced by all families when enrolling children irrespective of whether the tuition is free or not. In addition, the data do not provide a measure on income to be used as a proxy for full income, but has a measure of consumption. Figure 3.3(a) shows a histogram of full income of families, with values ranging from c1,590 to c77,905. The second graph figure 3.3(b) shows the school costs used in the estimation, which range from c12 to c1438. In addition to variation in school costs and consumption, the level of the



(a) Household Consumption



(b) School Costs

Figure 3.3: Data Variation

school grant also varies depending on the number of children in the household. The treatment effect is estimated by matching untreated individuals with other untreated individuals on functions of observable characteristics. Identifying the *ex ante* treatment effect also requires that any unobserved heterogeneity (ν) remains the same before and after treatment ie. (ν) is independent of consumption and school costs. However, the use of consumption in the reduced form means there is the problem of potential endogeneity. This would particularly be the case if decisions about schooling depend on the opportunity costs of enrolment or if parents are heterogeneous in preferences to invest in children's education. These preferences are likely to depend a great deal on family structure and background etc. In the case of preventive care utilization, consumption could be endogenous if accessing these services affects work decision/earnings. To make this assumption plausible, empirically the matching functions include a set of family characteristics - X_h and price variables.

$$f(\nu|F, \theta, X_h) = f(\nu|\tilde{F}, \tilde{\theta}, X_h)$$

3.4 Empirical Specification

The behavioural model reduced form approach proposed by Todd & Wolpin (Forthcoming) exploits exogenous variation in the policy related variables (in this paper full income and school costs) at the baseline and matches untreated individuals on functions of observable characteristics. The estimator they propose is broadly analogous to an *ex post* matching estimator set in the potential outcomes framework (Neyman 1990, Rubin 1974). An *ex post* matching evaluation uses information on both treated outcomes (M_1) and untreated outcomes (M_0) from a suitable comparison group and matches individuals on observable characteristics where the average outcomes for the matched untreated individuals is the counterfactual for the average outcomes for the treated group if they had not been treated (Heckman, Ichimura & Todd 1997). In the case of the *ex ante* evaluation the outcomes of the treated group (M_1) are unobserved and must be estimated from the observed (M_0) untreated/baseline information. This translates to estimating:

$$M_{1i} = E[M_{0j}|F_i = \tilde{F}_j, \theta_i = \tilde{\theta}_j, p_{mi} = p_{mj}, n_i = n_j, g_i = g_j, X_{hi} = X_{hj}] + \epsilon \quad (3.10)$$

The estimator proposed by Todd & Wolpin (Forthcoming) for the *ex ante* evaluation estimates the average treatment effect for those eligible for the program (intent-to-treat (ITT)) as:

$$\alpha = \frac{1}{k} \sum_{j=1}^k \sum_{i \in S_m} E(M_i | F_i = \widetilde{F}_j, \theta_i = \widetilde{\theta}_j, p_{mi} = p_{mj}, n_i = n_j, g_i = g_j, X_{hi} = X_{hj}) - M_j(F_j, \theta_j, p_{mj}, n_j, g_j, X_{hj}) \quad (3.11)$$

3.4.1 Estimating School Costs

Implementing the above matching estimator requires estimation of the unobserved treated outcomes as a function of consumption, school costs and a set of household characteristics. School costs (δ_i) are determined by the enrolment status of the child and hence are observed in the data for only those children who are currently enrolled in school and zero costs observed for those not enrolled. The problem of predicting school costs for the entire sample of children requires using a two-step process decomposing the participation decision and the determinants of the cost of schooling. A *two-part model (2PM)*¹ is applied where in the first part, the enrolment decision, is modelled using a probit and the second part predicts the cost of schooling as a linear function of the determinants of school costs (Mullahy 1998). The most common specification of the second part is a log transformation of the outcome variable. A problem with using a retransformed OLS in this case is that zero school costs are also observed in the sample of those children currently attending school. A log transformation would drop these observations from the estimation sample. A further problem arises with retransformation of the outcome variable to the original scale in the presence of heteroskedasticity. Manning (1998) shows that heteroskedasticity leads to biased estimates of the outcome variable and correction requires determining whether the heteroskedasticity is across different groups or caused by a particular subset of the covariates. To overcome these issues the second part of the *2PM* is estimated using the *extended estimating equations model*

¹A bivariate selection model was initially estimated and the non-linearity of the inverse mills ratio showed no evidence of selection bias. Also, school costs are not normally distributed and a log transformation would drop the observations that indicated zero costs. A further difficulty arises in finding a suitable exclusion restriction that affects school enrolment but not school costs. To overcome these problems a two-part model is used.

(*EEE*) proposed by Basu & Rathouz (2005). The *EEE* approach is a semi-parametric extension of a standard *generalized linear model (GLM)* incorporating flexible link and variance functions. It has two main advantages it identifies an appropriate link function from the data enabling identification of an underlying model for the error distribution and when no particular distribution can be identified from the outcome variable it serves as a robust estimator. Specifically, the *EEE* combines a Box-Cox transformation for the link function and includes a class of link functions represented by an estimated parameter λ :

$$\begin{aligned} \frac{\delta^\lambda - 1}{\lambda} & \text{ if } \lambda \neq 0 \\ \log(\delta) & \text{ if } \lambda = 0 \end{aligned}$$

It also allows for heteroskedasticity and uses a general power function for the variance defined by two-parameters θ_1 and θ_2 :

$$V(y) = \theta_1 \delta^{\theta_2}$$

The model is estimated separately for boys and girls.

3.4.2 Estimating Counterfactual Outcomes

In this paper the ex ante impact on two binary preventive care measures are estimated - health checks for children under 3 years of age and full coverage of vaccinations for children between the ages of 12-23 months. The unobserved binary outcomes $E(M_i|F_i = \tilde{F}_j, \theta_i = \tilde{\theta}_j, p_{mi} = p_{mj}, n_i = n_j, g_i = g_j, X_{hi} = X_{hj})$ are estimated using Klein and Spady's (Klein & Spady 1993) semiparametric estimator. The estimator belongs to the class of non-parametric single index models with conditional probability $P(M = 1|X = x) = G(x\beta)$ but where the distribution function G is left unspecified. The estimator is semiparametric in the sense that the only nonparametric component of the estimator is the linear index $G(x\beta)$ while the x 's maintain a linear specification as in the parametric counterpart of a probit or logit. The semiparametric single index specification offers several advantages over a fully nonparametric approach, it allows for as many covariates as required by the model by eliminating the *curse of dimensionality* problem where the model's convergence rates are inversely proportionate to the number of covariates. It also offers greater regions of support for predic-

tions as compared to a nonparametric model by extending the region of support beyond the observed x to points not in the support of x but in the support of $x\beta$ (Horowitz 1998).

Klein and Spady adapted single index models for binary outcomes, the index function is defined as:

$$E(M|x) = P(M = 1|x) = G(x\beta)$$

The estimator, like more general single index models, involves two unknowns - β and G . Estimation of both elements require several identification restrictions. Similar to all linear models, identification of β requires G to be a non-constant function along with the absence of multicollinearity amongst the covariates. In addition, to uniquely identify the function $G(x\beta)$ single-index models involve *location normalization* and *scale normalization* restrictions. Location normalization is achieved by requiring the covariate vector to include no intercept term while scale normalization involves restricting the β coefficient of one continuous variable to equal one. Identification in single-index models is achieved because the conditional mean function can remain constant with changes in x as long as the index $x\beta$ remains constant. However, with continuous covariates a constant index (ie. $x\beta = k$) for a given set of covariates has probability zero. To overcome this a further identification restriction is required where G is a differentiable function so that $G(x\beta)$ is close to $G(k)$ when $x\beta$ is close to k (Horowitz 1998). A final set of restrictions are required when X contains both discrete and continuous variables. The first of these requires that the discrete elements of the covariate vector do not divide the support of $x\beta$ into disjoint subsets. The final restriction is referred to as the 'non-periodicity condition' for the function G requiring it to be strictly increasing.

Klein and Spady's adaptation of single index models for binary outcomes uses maximum likelihood estimation(MLE) where the log-likelihood is:

$$\ln L(\beta, G_n) = n^{-1} \sum_{i=1}^n [M_i \ln G_n(x_i\beta) + (1 - M_i) \ln(1 - G_n(x_i\beta))] \quad (3.12)$$

The difference from a parametric estimator such as probit or logit is that $G_n(x_i\beta)$ is a semi-parametric likelihood estimate which is estimated using a leave-one-out nonparametric esti-

mator of the density of $x\hat{\beta}$ conditional on M , where for any z

$$G_n(x_i\beta) = \frac{P_n g_n(z|M=1)}{P_n g_n(z|M=1) + (1 - P_n) g_n(z|M=0)} \quad (3.13)$$

where g_n is the kernel estimate of the conditional density of $x\beta$ ($g(\cdot|M)$) and g_n is defined as:

$$g_n(z|M=1) = \frac{\sum_{i=1}^n M_i K(z - x_i\hat{\beta})/h_n}{nP_n h_n} \quad (3.14)$$

$$g_n(z|M=0) = \frac{\sum_{i=1}^n (1 - M_i) K(z - x_i\hat{\beta})/h_n}{n(1 - P_n) h_n} \quad (3.15)$$

where P_n is the empirical probability $P_n = \sum_{i=1}^n M_i$, the proportion of infants taken to a health check in the last six months or the proportion with full coverage immunization. K is a kernel function and h_n is the bandwidth.

Klein and Spady show that the estimator is asymptotically efficient and achieves the semi-parametric efficiency bounds of Chamberlain (1986) and Cosslett (1987). The resulting vector of parameter estimates ($\hat{\beta}$) is shown to have the following properties:

$$n^{1/2}(\hat{\beta} - \beta) \rightarrow_d N(0, \Omega)$$

$$\Omega = E \left\{ \left[\frac{\partial G(X_i\beta)}{\partial \beta} \right] \left[\frac{\partial G(X_i\beta)}{\partial \beta} \right]^T \left[\frac{1}{G(X_i\beta)(1 - G(X_i\beta))} \right] \right\}^{-1}$$

$E(M_i|F_i = \tilde{F}_j, \theta_i = \tilde{\theta}_j, p_{mi} = p_{mj}, n_i = n_j, g_i = g_j, X_{hi} = X_{hj})$ (unobserved treated outcomes for health checks and FCV) are estimated using the semiparametric estimator described above. Each of the unobserved outcomes is estimated using a different model specification. They include the basic household characteristics - education and gender of the household head, composition of the household, school costs and full wealth and age of the child. In addition the model for health checks includes two community level characteristics that serve as opportunity costs of accessing care - distance to the nearest health facility and distance to the nearest public transport facility. In estimating the model for full coverage of vaccination predicted outcomes are compared with Barham & Maluccio (2009) who estimate two versions of the experimental outcome - one including community variables and the other without. They find both give similar results. The estimation in this paper is a model without community level characteristics and hence includes only the household and child characteristics described above.

Scale normalization is achieved by setting the coefficient for years of education for the household head equal to 1. The estimated models are then used to first predict outcomes for the observed control (baseline) group observations and then extrapolate the predictions under treatment by evaluating the function at $(\tilde{F}, \tilde{\theta})$.²

The above specifications are also used for simulating alternate policy scenarios for both the health outcomes and school enrolment. For the simulation of school enrolment the model consists of the same household and child characteristics described earlier and include variables for distance to the nearest primary and secondary schools, whether the household lives in a coffee growing community and distance to the nearest public transport facility. The enrolment models are estimated separately for boys and girls and jointly for different age groups.

Both within sample predictions and extrapolation can only be carried out in regions of common support. In the original formulation of the model Klein and Spady trim the likelihood function (equation 3.12) to ensure that G is bounded away from 0 and 1. Simulations of

²The statistical package np (Hayfield & Racine 2008) available for the software R was used. The scalar bandwidth for the index $x\beta$ for the health checks model is 0.037 and for the FCV model is 0.077.

their model and other studies using this estimator (Horowitz 1993, Gerfin 1996, Fernández & Rodríguez-Poo 1997) find little impact of trimming in empirical applications. Following these studies this application does not trim the likelihood function. Extrapolation in nonparametric models is only valid at points with positive data density; hence while not trimming the likelihood function, trimming is carried out to define the region of common support i.e to identify regions of positive data density in the extrapolated values. The region of support S_m in the semiparametric model is defined as $S_m = \{x\beta \in R^2 \text{ such that } f(x\beta) \geq 0\}$ where $f(x\beta)$ is the nonparametric density of the linear index.³ Heckman, Ichimura & Todd (1997) propose that the density should be strictly positive as defined by S_p and should exceed a minimum cut-off to avoid points with very low density. Thus the extrapolation is valid for only those points of evaluation where

$$f(x\hat{\beta}) > c \tag{3.16}$$

Heckman, Ichimura & Todd (1997) recommend setting the cut-off at a percent quantile of the estimated densities. Here c is set at the 2% quantile. Only those observations that meet the above criterion are kept in the extrapolation sample.

3.5 Data and Variables

The ex ante evaluation is carried out using data from the randomized pilot evaluation of RPS. As part of the experiment data was collected in four rounds. The first, a census survey of chosen localities in May/June 2000, this was followed by a comprehensive baseline for all members of the treatment and control groups, followed by follow-up surveys in October 2001 and October 2002. This ex ante evaluation uses data from the census and baseline surveys for the estimation and compares the predicted outcomes with those observed using the first follow-up survey. A total of 1581 households were included in the experiment.

The estimation sample consists of 576 households with children aged 0-3 who were eligible for the food security, health and nutrition component of the program. Of this sample 353

³The densities are estimated using the method of Li & Racine (2003) who use 'generalized product kernels' for mixed data. The bandwidths were set using maximum likelihood cross validation

households also had children between ages 7-13 who had not completed grade 4 of primary school and were hence eligible for the school transfer as well.

The census survey collected basic information on living conditions and the distance to the nearest school for each household and at individual level, information on education and school enrolment, land ownership and work. The baseline and follow-up surveys collected detailed information on the same categories and in addition the baseline survey had a detailed module on health and health related variables for all members of the household. These surveys however did not gather information on income and different sources of income but instead collected information on complete household expenditure and asset ownership.

3.5.1 Dependent variables

The ex ante impact is estimated for two health related variables - health checks and full coverage of vaccinations (FCV). In the randomized experiment parents of all children below 5 years in the household were asked whether the child had been taken for a health check in the last six months. The estimation of predicted impact focuses on children below 3 years with a balanced (households observed in the baseline and follow-up survey) sample size of 792. At the baseline just over 70 percent of children in this age group had been taken to a health check within the last 6 months.

The full coverage of vaccination variable was constructed following the approach in Barham & Maluccio (2009), using a series of questions on different vaccines. The baseline survey recorded the number of doses each child received since birth for the following set of vaccines (1) tuberculosis vaccine (BCG) (2) measles containing vaccine (MCV) or measles-mumps-rubella (MMR) vaccine (3) oral polio vaccine (OPV) (4) diphtheria-pertussis-tetanus vaccine (DPT) or pentavalent vaccine (or both). The international recommendation for up-to-date time of vaccination is <12 months and 12-23 months of age depending on the vaccine. A child is considered fully vaccinated if it receives all the required doses within the specified time period. Table 3.1 gives the schedule for recommended vaccinations. A binary variable for FCV was constructed equal to one if a child's vaccine schedule was up-to-date and zero

otherwise. The estimation was restricted to children 12-23 months of age to be sure that all children had a chance to receive the BCG vaccine which is scheduled for below 12 months. The total sample size of the balanced sample is 281.

Table 3.1: Basic vaccination schedule for up-to-date vaccinations

Disease	Vaccine	Dose	Recommended age
Tuberculosis	BCG	1	At birth
Measles	MCV	1	12 months
Polio	OPV	3	2, 4, 6 months
Diphtheria-pertussis-tetanus	DPT	3	2, 4, 6 months

Source: Barham & Maluccio (2009)

A second objective of this paper is to evaluate alternate policy scenarios. These simulations are carried out for both the schooling outcomes as well as the two health related variables described above. The main policy variable considered for the school outcome is the enrolment rate. Households with children between ages 7-13 who had not completed grade 4 of primary school were eligible for the school transfer conditional on all children in the household enrolling and maintaining 85% attendance. The enrollment outcome is also a binary variable equal to one if an eligible child enrolls and zero otherwise. The estimation sample consists of 1786 program eligible children.

3.5.2 Explanatory variables

The estimation of the unobserved outcomes under treatment is driven by the reduced form equations from the theoretical model. The policy enters the model through an impact on the budget constraint and changes two key variables in the reduced form - school costs (for families with children eligible for the school transfer) and full wealth (for all families). The baseline surveys provide information on total household expenditure and asset ownership which are together used as a proxy for full wealth. The second policy variable - school costs is however observed only for those children aged 7-13 who are currently enrolled in school and have to be estimated for the those who are eligible but not enrolled. For those families in the sample without children eligible for the school component school costs remain zero.

The schools costs are estimated using variables that capture direct and opportunity costs, set of household and child characteristics and family wealth. The census survey provides

information on the education of the household head and distance to the nearest primary school (used as a measure of opportunity cost of travel time). These variables are mapped to the baseline survey. The baseline survey provides all the other variables. Family characteristics include - household expenditure and asset ownership, age, gender and years of schooling completed by the household head, number of children of school going age, number of adults in the household and number of children under 5 years. To avoid problems of endogeneity, household wealth (expenditure plus assets) is included net of the school costs and health expenditure.

The unobserved outcomes $E(M_i|F_i = \tilde{F}_j, \theta_i = \tilde{\theta}_j, p_{mi} = p_{mj}, n_i = n_j, g_i = g_j, X_{hi} = X_{hj})$ for health checks and FCV are estimated using the variables generated by the reduced form of the behavioural model. These include, household characteristics - education and gender of the household head, composition of the household, school costs and full wealth and age of the child, distance to the nearest health facility and distance to the nearest public transport facility (for the health checks model). The simulations of alternate policy scenarios also use the same specifications.

3.6 Results

3.6.1 Estimating School Costs

Table 3.2 shows the results from estimating the two part model for boys and girls. The probit participation model for both boys (1) and girls (3) show a similar pattern, with enrolment being most likely between the ages of 8 and 10 as compared to children aged 7 (reference category) and declining with older children. Boys drop out earlier (above age 10) while girls aged 13 are less likely to enrol when compared to the reference group. This pattern follows most developing countries where many children enrol and stay in school only for a few years, dropping out between the ages of 11-13 to find employment. Consumption net of school costs and education of the household head are significant and have a positive impact on enrolment. As mentioned earlier the probit model includes the number of children under 5 years as a proxy for child labour. The estimates show similar negative magnitudes for boys and girls

indicating having younger children in the household decreases the likelihood of enrolment. A similar effect of distance to the nearest school is observed, with children being less likely to enrol if schools are further away. Enrolment probabilities differ for boys and girls depending on the gender and the employment status of the head of the household. Girls are less likely to enrol if a male is head of the household, as is the case in 88% of the households in the sample. The direction of the coefficient for employment status is less intuitive as boys seem less likely to enrol if the household head is employed. This result is probably due to the nature of employment, with about 85% of the sample being involved in farm activities. The last two variables though not significant in the model do indicate the presence of a gender gap from additional opportunity costs for boys and cultural differences that contribute to the differences in schooling.

Columns (2) and (4) of Table 3.2 provide results from the second part of the two part model using the *extended estimating equations model (EEE)* (Basu & Rathouz 2005) for school costs.⁴ Boys in the reference category (age 7) face the highest school costs. At other ages there is no significant impact on school costs. For girls however, school costs increase with age. Families with greater consumption tend to spend more on education, although more on the boys than the girls. In both cases children of the same age and children under five is significant (except for girls -children under5) and negative. This is intuitive in the sense that sharing of resources reduces the costs per child as the number of school age children increases.

In Column (2) for the boys sample the link parameter is estimated to be $\lambda = 0.289$ (95% C.I: 0.01, 0.57). The variance function represented by $\theta_1 = 1.2$ (95% C.I:1.07 ,1.42) and $\theta_2 = 1.6$ (95% C.I:1.39 , 1.80) is close to a gamma distribution. Column (4) provides the estimates for the sample of girls. In this case with $\lambda = 0.66$ (95% C.I: 0.26, 1.06), the link function is close to a square root link. The values $\theta_1 = 1.5$ (95% C.I:1.26 ,1.87) and $\theta_2 = 1.74$ (95% C.I:1.51, 1.95) again suggest a gamma distribution.

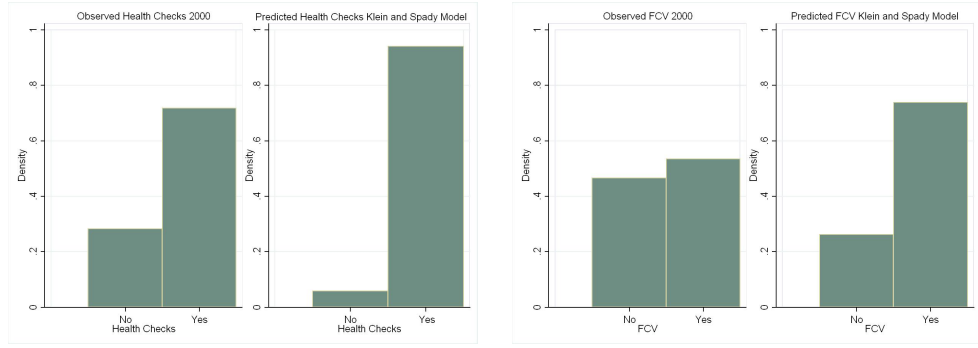
⁴An alternative approach to the EEE model would be to use a *generalized linear model* with a specified link function and distribution. However, failure to specify the correct link function results in misspecification of the model. To avoid such misspecifications, the EEE approach was used since it does not require an *a priori* assumption of a link function or distribution. This approach 'helps to identify an appropriate link function and to suggest an underlying distribution for a specific application but also serves as a robust estimator when no specific distribution for the outcome measure can be identified' Basu & Rathouz (2005).

Table 3.2: Estimating School Costs

VARIABLES	(1) Probit-Boys Enrollment	(2) EEE-Boys School Costs	(3) Probit-Girls Enrollment	(4) EEE-Girls School Costs
age8	0.116 (0.151)	0.117 (0.139)	0.251 (0.160)	0.168 (0.103)
age9	0.265 (0.161)	0.180 (0.147)	0.544** (0.168)	0.385*** (0.103)
age10	0.174 (0.164)	0.0439 (0.124)	0.285 (0.172)	0.345* (0.172)
age11	-0.00216 (0.164)	0.0458 (0.133)	0.202 (0.179)	0.470*** (0.116)
age12	-0.00164 (0.172)	0.189 (0.137)	0.147 (0.188)	0.523** (0.197)
age13	-0.554*** (0.168)	0.0256 (0.138)	-0.159 (0.198)	0.205 (0.134)
HH Cons (adjusted)	0.0000116** (0.00000447)	0.0000405*** (0.00000345)	0.0000122* (0.00000593)	0.0000264*** (0.00000359)
School dist	-0.00703*** (0.00178)	0.00346** (0.00118)	-0.00806*** (0.00173)	0.00230 (0.00155)
No. of adults	-0.0336 (0.0328)		0.0536 (0.0413)	
Children under5	-0.172** (0.0550)	-0.224*** (0.0398)	-0.228*** (0.0621)	-0.0780 (0.0430)
Children 7-13	0.0814 (0.0463)	-0.258*** (0.0443)	-0.0462 (0.0557)	-0.240*** (0.0391)
HHH gender	0.335 (0.183)		-0.0755 (0.198)	
HHH age	0.00714 (0.00529)		0.00979 (0.00562)	
HHH yrs of ed	0.106** (0.0381)		0.180*** (0.0436)	
HHH works	-0.0949 (0.173)		0.186 (0.194)	
Constant	-0.0227 (0.325)	-0.377* (0.155)	-0.0752 (0.375)	-0.288* (0.131)
λ		0.289* (0.143)		0.663** (0.204)
θ_1		1.242*** (0.0887)		1.564*** (0.157)
θ_2		1.597*** (0.106)		1.737*** (0.111)
Observations	945	687	845	631

Robust standard errors in parentheses, clustered at the household level

*** p<0.01, ** p<0.05, * p<0.1



(a) Observed Baseline (2000) Outcomes and Predicted Outcomes for Health Checks (b) Observed Baseline (2000) Outcomes and Predicted Outcomes for FCV

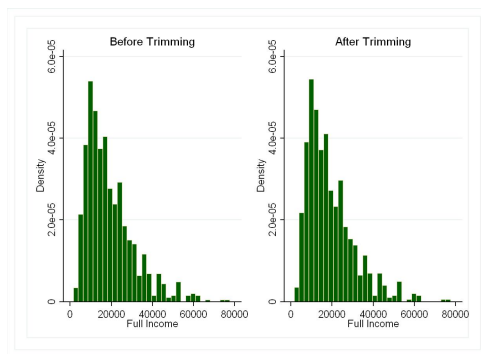
Figure 3.4: Comparing Observed and Predicted Outcomes

3.6.2 Predicting Impact

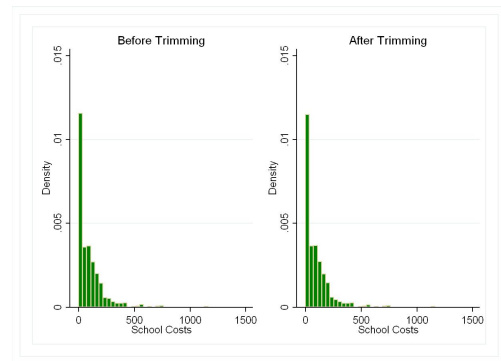
The counterfactual (unobserved) health related outcomes after program implementation ($M^{**} = \Psi(\tilde{F}, \tilde{\theta}, p_m, n, g; X_h)$) are estimated using the Klein and Spady binary estimator described in section 3.4.2. The health checks model is specified as a function of education and gender of the household head, composition of the household, estimated school costs and consumption (net of education and health expenditure) of the household after adjusting for the income transfer, and age of the child. Figure 3.4(a) compares the observed baseline outcomes (M_j in equation 3.11) with the predicted impact from the Klein and Spady model (M_i in equation 3.11). The figure shows a large increase in the estimated proportion of children being taken to health checks with the cash transfer program.

Figure 3.4(b) shows similar graphs for the FCV model. When compared to the baseline, the cash transfer induces a significant increase in the proportion of children under 2 years who have up-to-date immunizations.

The predictions from the estimator (equation 3.11) are only valid in the region of common support defined by equation 3.16. Figures 3.5(a),3.5(b),3.5(c)and 3.5(d) compare the distributions of the policy variables before and after trimming at 2% quantile of the density of the linear index ($f(x\hat{\beta})$). The graphs show that observations with very low density - concentrated at the extreme right tail of the distributions are eliminated. Intuitively this means

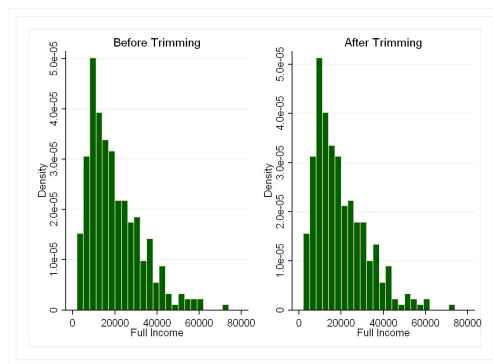


(a) Household Consumption

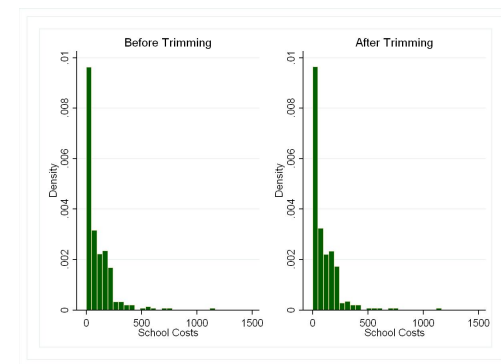


(b) School Costs

Health Checks Estimation



(c) Household Consumption



(d) School Costs

FCV Estimation

Figure 3.5: Trimming Klein and Spady estimations

that families with very high consumption or school costs where no suitable matches could be found are dropped from the estimated treatment effect.

Conditional cash transfers provide a social policy tool that combine improvement in child health with education outcomes for school eligible children. The cash transfers are typically accompanied by a package of health related conditionalities relating to preventive care. But combining both does not allow separation of different pathways ie. the impact of the cash transfer separate from the conditionality. The RPS program was also designed as a combination of three factors - a school enrolment transfer, a food security transfer and a series of conditionalities relating to child health. But as described in the program description the RPS program for the first 8 months of the program health services were not provided nor any conditionalities enforced relating to the health/food security transfer. The impacts for the first year are then largely the impact of the cash transfer in improving utilization of health services without the conditionalities. However, it is possible that the last four months of implementing the conditionalities did have some effect on the experimental outcomes. The double difference estimates presented in 3.3 could contain the combined policy effects from the last four months. In addition during the first year of the program there was also an improvement in outcomes in the control group. One obvious reason for this could be positive spillovers from the program to the control group â Barham & Maluccio (2009) test for this and find no evidence of spillovers. Alternatively, due to the general strengthening of services in the study area these could have benefited control groups as well. A similar rise was not observed in non-study areas of Nicaragua (Barham & Maluccio 2009). As a result the experimental double difference benchmark impacts are conservative and likely downward biased.

Table 3.3: Predicted Impact

OUTCOMES	(1) Predicted Impact	(2) Sample sizes [@]	(3) Experimental impact
Health Checks (children below 3 years)	0.22*** (0.0250)	779 / 792	0.24*** (1 year)
FCV (Children below 2 years)	0.20*** (0.0451)	275 / 281	0.20**(1 year)

@ treatment observations after trimming, total number of observations.

Bootstrapped standard errors clustered at the comarca level (500 reps).

*** p<0.01, ** p<0.05, * p<0.1

Table 3.3 shows the one year predicted impacts for the two preventive healthcare utilization

outcomes and compares them with the double difference estimates from the experimental evaluation. Column (1) reports the results from the Klein and Spady estimates, column (2) provides the sample size before and after trimming the predicted outcomes and column (3) gives the experimental estimates. The impacts in this table are for the balanced sample defined as observing a household in both the baseline and follow-up survey. The *ex ante* model predicts closely the outcome of health checks in the last six months for children below 3 years, with one year of cash transfers resulting in a .22 increase as compared to a .24 increase in the experimental evaluation. As mentioned earlier a rise of 3 percentage points was also seen in the control group that could not be explained by spillovers but is most likely due to the improvements made in the program region by the Government. Given this improvement, the *ex ante* results already underpredict (marginally) the experimental effect on health checks and are likely to be the true unconditional effects. The table also reports the one year impact of on-time full coverage of immunization for children between 12 and 23 months of age. The *ex ante* result shows a statistically significant .20 increase in FCV which perfectly predicts the results from the experimental evaluation. In the case of FCV the control group had a 10 percentage point increase which shows that the *ex ante* results are much lower. These results demonstrate that it may be possible to improve utilization of preventive services in low income households without implementing conditionalities on their usage. Of particular interest is the immunization outcome which in the RPS design is not a pre-requisite for receiving the cash transfer. An improvement in household income encourages households to invest in child health. Other studies have found similar results of cash transfers or income increases leading to improvements in child health. Fernald et al. (2008) disaggregate the *ex post* impact of cash transfers in Mexico's Oportunidades (Progresa) program from other aspects of the program. They hold constant the conditionalities as required by the program and analyse the impact of increases in the cash transfer. Since in Progresa the conditionalities for preventive health were enforced rigorously from the beginning they look at direct health outcomes such as stunting and find improvements in child health from larger transfers.⁵

Comparing the *ex ante* results to the experiment provides a way of validating the model used. The validated models are then used to simulate alternate policy scenarios. Two alternate policy formulations are estimated using the validated models, the first estimates the impact of reducing the total cash transfer amount to 75% of the original program while continuing to

⁵Case et al. (2002) also find evidence that children in families with higher income have better health outcomes.

Table 3.4: Simulating Counterfactual Policy Scenarios

OUTCOMES	(1) Experimental Impact	(2) 75% of origi- nal	(3) Unconditional transfer (food security)	(4) Sample sizes [@]
Health related outcomes				
Health Checks	0.24***	0.22*** (0.0264)	0.21*** (0.0277)	779 / 792
FCV	0.20***	0.21*** (0.0431)	0.12*** (0.0427)	275 / 281
School enrolment outcomes				
Boys 7 -13	0.19***	0.19*** (0.0207)	0.17*** (0.0209)	859 / 876
Girls 7 -13	0.20***	0.21*** (0.0224)	0.16*** (0.0169)	754 / 767
Boys & Girls <10	0.23***	0.16*** (0.0239)	0.12*** (0.0263)	829 / 844
Boys & Girls >=10	0.15***	0.14*** (0.0153)	0.17*** (0.0123)	786 / 799

@ treatment observations after trimming, total number of observations.

Bootstrapped standard errors clustered at the comarca level (500 reps).

*** p<0.01, ** p<0.05, * p<0.1

maintain the conditionalities. Both cash transfers, the school transfer and food security/health transfer are reduced to 75% of the original amount. The per child component of the school transfer is also reduced similarly. The results of this scenario are reported in column (2) of Table 3.4. The second scenario estimated is that of providing a transfer equal to the value of the food transfer. This component has two aspects, first the conditionality on enrolment is removed ie. school costs are no longer adjusted but secondly, the conditionality for receipt of the transfer ie. mothers attending health workshops continues. This is how the original program was modeled except that there is no school component of the transfer. As mentioned earlier the conditionalities on accessing preventive care were not enforced in the first 8 months. The estimated impacts can provide an insight into how households are likely to allocate resources between preventive health care for young children and education of older children. The results of the second simulation exercise are reported in column (3) of Table 3.4. In both cases the estimated results are compared to the experimental results (column 1) from the original program design. The last column of Table 3.4 reports the sample sizes for the models before and after trimming.

Comparing column (1) and column (2) of Table 3.4 shows that reducing the amount of the transfer to 75% of the original amount has little impact on the preventive care utilization outcomes. Health checks increase by .22 as compared to .24 in the original program while FCV shows a 1 percentage point difference from the experimental impact. The price effect of the school transfer is equally strong when reduced to 75%. Overall boys and girls show similar outcomes from the reduced cash transfer scenario as they did under the original program. For boys aged 7-13 both the original program and the reduced cash transfer show a .19 increase in school enrolment. The simulation for the sample of girls shows a 1 percentage point increase difference of .20 to .21. Similar results are evident when looking at enrolment by age groups. The same specification was used for the models for boys and girls and the different age groups. A point of concern here is the predicted impact for boys and girls below 10. The predicted impact under the alternate policy scenario shows only a .16 increase in enrolment as compared to .23 from the experimental outcome. Some caution is required in interpreting this result as the *ex ante* model of the original program underpredicts the impact.⁶ The estimated impact for the reduced cash transfer is very close to the estimated *ex ante* result from the original policy design. In general a reduction to 75% while maintaining the conditionality for the school component has important budget implications for the program. The price effect from reducing school costs by conditioning enrolment can be achieved with a lower level of transfer. Even at the lower level of the transfer families access preventive care at the same rate as the original program specification.

To test the importance of the conditionality of the school transfer and its implications for other components of the household's behaviour the validated model is used to estimate the impact of a transfer equal to just the food security/health component with no adjustment of school costs, thus relaxing the enrolment conditionality. Column (3) in Table 3.4 reports the estimated impacts which are quite different from both the experimental results and the policy scenario with 75% of the transfers and the conditionalities. School enrolment for boys is lower by 2 percentage points while for girls there is a 4 percentage point difference. More interestingly there is a large reduction in enrolment for children below 10 (.12 increase from the baseline) and a rise in enrolment for children above age 9 years (2 percentage point increase). The biggest difference is in the proportion of children who receive on-time full

⁶The *ex ante* estimation of the original program shows a statistically significant .17 increase in enrolment for children below 10 years. Results for all the other age groups are statistically significant and almost identical to the experiment: boys .19, girls .21, children 10 years and over .15.

coverage of vaccinations, the unconditional transfer results only in a .12 increase from the baseline while taking children for health checks shows a 3 percentage point decline when compared to the experimental impact. The less obvious income effect for health checks could be the nature of the question in the survey which refers to a very short term of a health check in the last 6 months as compared to the repeated visits over a longer time frame as required by the vaccination outcome. One plausible explanation for the sharp decline in enrolment amongst younger children and in the health outcomes is that without a compulsory enrolment requirement for all children aged 7-13, parents now have to choose between enrolling older or younger children. The school sample consists of children who have not completed grade 4 of primary school. In the short term parents seem to invest in the enrolment of the older children while delaying enrolment for the younger children. Also, the cash transfers are meant to substitute for the wages earned by the children enabling the family to have them enrol in school. In the short term with a fixed budget constraint parents seem to compromise accessing preventive care, particularly ensuring up-to-date vaccinations for infants, and enrolling younger children for an extra year of education for older children who are less likely to complete further education if they delay enrolment further. The idea of the food transfer in addition to providing a school transfer was to improve nutrition and remove financial barriers to accessing health care. This however does not appear to be the pathway with just a cash transfer with no conditionality on schooling. Particularly for immunizations a strong income effect emerges with greater income levels resulting in higher rates of child immunizations. However, to improve investment in child health and school enrolment the conditionalities seem critical, particularly to ensure enrolment of younger children is not compromised for education of older children.⁷ Todd & Wolpin (Forthcoming) also find for the Progres program that conditionalities are necessary to have an impact on enrolment, although in their sample they study older children (ages 12-15).

3.7 Conclusion

This paper applies the methods proposed by Todd & Wolpin (Forthcoming) to estimate *ex ante*, the impact of Nicaragua's conditional cash transfer program Red de Protección So-

⁷Simulations with just the school enrolment component - with conditionalities shows schooling results very close to the those of *ex ante* model for the original program specification and also shows evidence of an income effect on FCV.

cial. This approach relies on reduced form estimation of the impact by matching untreated individuals on functions of observable characteristics. Identification of the impact relies on variation in the variables related to the policy being evaluated, in this case school costs and full wealth. A simple household model of school enrolment and investment in child health is used to derive the reduced forms that are empirically estimated. The approach is used to estimate the impact of the cash transfer on accessing preventive care for children ie. taking children below 3 years to health checks and full coverage of vaccinations for children between 12-23 months. Empirically the model is implemented using a semi-parametric single index framework that allows for an increase in the dimensionality of the covariate vector. The outcomes are binary and the semi-parametric estimator proposed by Klein and Spady is used to predict the unobserved outcomes under treatment. The data set combines baseline data from the RPS experiment along with some information from the census survey . The baseline data is used as a single cross-section combining both control and treatment groups. The estimated *ex ante* impact is validated against results from the randomized experiment. These models are then used to simulate two policy scenarios that differ from the original program in the amount of the transfer and the conditionality. These simulations are carried out for both school enrolment (children aged 7-13 years) and preventive health care utilization for younger children.

In general the estimated impacts all have the same direction as the experiment. The model performs well in predicting the magnitude of the impact for the two health related outcomes. An improvement in household income increases a households' investment in child health. The two policy simulations show interesting results. The first simulation maintains the school related conditionality of the program but reduces both cash transfer components by 75%. The estimations show little difference in most of the categories from the original program. The price effect from reducing school costs by conditioning enrolment can be achieved with a lower level of transfer as can similar levels of preventive care utilization. The second simulation of a cash transfer equal to the amount of just the food security component without the school related conditionalities shows a strong income effect for immunizations and is accompanied by change in enrolment patterns, with higher enrolment levels for older children but at the cost of delays in enrolling younger children.

Chapter 4

Ex ante Forecasting of the Distributional Effects of a Conditional Cash Transfer Program on Child Health in Nicaragua

4.1 Introduction

The impact of conditional cash transfer programs on child health is well documented in the literature (Gertler 2004, Fernald et al. 2008, Maluccio & Flores 2005). However these studies restrict themselves to the average impact of the programs and do not highlight distributional consequences. Even in cases where limited impact is found Heckman, Smith and Clements (1997) emphasize the importance of extending program evaluations to explore beyond average impacts. The heterogeneity can exist either across different observed covariates such as gender and age or the treatment effect itself may not be the same across all individuals. Following the paper by Djebbari and Smith (2008) in which the authors provide a formal analysis of heterogeneity in treatment effects, using data from Mexico's social protection program Progresa, there has been a small but growing literature on estimating distributional impacts in evaluating social programs in developing countries, particularly conditional cash transfer programs. This paper adds to the literature on distributional impacts of conditional cash transfer (CCT) programs by analysing the impact of Nicaragua's CCT program Red de Protección Social on child health; but it is unique in that it attempts to forecast these impacts from a pre-program scenario and by doing so makes an important contribution to the limited literature on ex ante evaluations of social programs.

Ex ante evaluations of social programs attempt to forecast the outcomes of a program before it is implemented (Todd & Wolpin 2006, Attanasio et al. 2005, Todd & Wolpin Forthcoming). These papers all rely on household behavioural models to forecast school education outcomes for the Progresa program. But while the first two rely on dynamic structural estimation procedures the last paper proposes the use of simple non-parametric estimation of reduced forms derived from an underlying behavioural model with minimal assumptions of functional form. In this approach exogenous variation in the variables relating to the policy are exploited to predict unobserved outcomes after the program by using a form of matching estimator that matches untreated individuals with lower policy variables with other untreated individuals with higher magnitudes of policy variables mimicking a post treatment scenario. The current paper is based on this reduced form approach of Todd & Wolpin (Forthcoming) but while they focus on the mean impact this study extends their idea to forecast changes in the entire distribution of weight-for-age z scores (WAZ) of children below 5. The data comes from the randomized experiment of RPS and uses information for just the control group. The estimated outcomes are then compared with the experimental results from the program.

This paper also adds to the existing literature by applying recently developed estimators to forecast distributional impacts. It draws on three different estimators - two semiparametric estimators, the first based on linear quantile regression (Melly 2005) and the second using distributional regressions (Chernozhukov et al. 2009); while the third uses a completely non-parametric specification to estimate the unobserved distribution (Rothe 2010). In all three cases changes in different quantiles of the WAZ distribution under treatment are estimated and compared with the linear quantile regression estimates from the experiment. Each of the estimators is a two stage process involving first-step estimation of the observed conditional distribution function of WAZ given a set of covariates. The second stage of estimating the unobserved unconditional (marginal) distribution of WAZ under the program involves integrating over the set of covariates in the post-treatment scenario. The main difference in the estimators lies in the first stage, each of the estimators uses one of the three methods mentioned above to estimate the conditional distribution function. In addition, tests for stochastic dominance are carried out to compare the estimated counterfactual distribution with the observed untreated distribution.

4.2 Background

4.2.1 Malnutrition and CCTs

Undernutrition of children is still widely prevalent in many parts and is associated with about half of child deaths around the world. Malnourishment makes children more susceptible to infections and less likely to survive other childhood illnesses such as diarrhoea and malaria. In 2007, UNICEF estimated that one out of four children in developing countries is underweight. It also has severe long-term consequences on growth and development. Malnutrition at an early age results in reduced mental and physical development and these children lose more days to illness than children not faced by this risk, resulting in poor school enrolment rates and education outcomes culminating in poor productivity and earnings in adulthood (Glewwe et al. (2001), Grantham-McGregor et al. (2007), Mendez & Adair (1999)).

Nutrition levels in children are largely determined by the food consumed, exposure to illnesses, availability and access to medical treatment and household and community factors. Cash transfer programs aim at reducing financial barriers to improving food intake, accessing education and health care while simultaneously investing in some household factors such as education of mothers to invest in better nutrition levels for the family. The objective is to reduce poverty while improving investment in human capital.

The positive relationship between income and child nutritional status is well documented in the literature. Typically the literature explores a longer term outcome - height-for-age (HAZ) Z score. In this paper the interest is in the short term impact of a conditional cash transfer program on child nutrition and hence the outcome weight-for-age Z score (WAZ) is used as a measure of nutritional status. In the literature the income effect of a cash transfer is expected to operate through food availability, women's education and access to clean water and sanitation facilities. Analyses range from cross country surveys to forecasting reductions in child malnutrition based on income growth. Haddad et al. (2003) use household survey data from 12 countries and malnutrition rates from a cross-section of countries to examine the relationship between income and nutrition status. The authors find that raising income growth

to beyond historical income growth rates will not reduce malnutrition sufficiently to achieve the 2015 Millennium Development Goals on child malnutrition and conclude that income growth must be accompanied by nutrition policies. Other studies looking to forecast reductions in child malnutrition include Glewwe et al. (2004) and Edmonds (2004), both of which use data from 1993 and 1998 Vietnam Living Standards Measurement Survey to measure income contribution to reduction in child malnutrition in this period. However, both these papers conclude that income growth accounts for only a small fraction of the improvement in child health status. A study that takes the analysis of these two papers further and is closest in method to the ones applied in the current paper is O'Donnell et al. (2009). This study estimates the full marginal distribution (counterfactual) from a change in covariates (over time) and decomposes the observed difference in child height into the contribution of an income change and a shift in the returns to an improvement in nutrition ie. 'nutrition function'. The similarity lies in the approach used to estimate the counterfactual distribution, the objective here though is not to decompose the change in nutrition status but to forecast the impact of a conditional cash transfer program across the distribution of WAZ and compare results with those of the experiment.

4.2.2 The RPS Program

The data for this paper is taken from the RPS conditional cash transfer program introduced in 2000 by the Government of Nicaragua. It is particularly suited to the exercise of forecasting outcomes as it was implemented as a two year randomized social experiment which enables comparing methods from the different estimators applied in this paper with those of experiment, providing a suitable benchmark for the results. Like all cash transfer programs RPS is a demand side social protection initiative that aims at reducing financial and informational barriers to accessing education and health care. It aims at reducing poverty across generation by encouraging investment in human capital. Two rural districts of Central Nicaragua - Madriz and Matagalpa were selected as pilot areas based on their poverty levels and capacity to implement the program. According to the 1998 Living Standards Measurement Survey 48% of Nicaraguans were classified as poor. The randomized experiment was implemented in 42 *Comarcas* (administrative units within municipalities) within the two districts chosen based on a marginality index.

The program had two components - education and food security/health. Under each component families received cash transfers conditional on their fulfilling certain requirements. In the education component families with children between the ages of 7-13 who had not completed grade 4 of primary school were eligible for the transfers conditional on the eligible children enrolling and maintaining 85% attendance. Families received two transfers (Maluccio & Flores 2005) - a 'school attendance transfer' provided as fixed sum for all families equaling the Córdoba 2000 equivalent of US\$112 per year; and a per child 'school supplies transfer' of US\$5. If any of the children did not meet the conditionality the family failed to receive the lump sum transfer.

The food security/ health component of the program also involved a per family cash transfer of US\$224. To receive this transfer two main conditionalities had to be met - mother's of children under age 5 had to attend in alternative months health education workshops, and the children aged below 5 had to be taken to scheduled preventive health care appointments. Services at the appointments were offered free of charge and included growth and development monitoring, vaccinations, provisions of antiparasites, vitamins and supplements. Program implementers found that there were frequent delays in the delivery of vaccines during these health checks and consequently a sub-conditionality of maintaining up-to-date vaccination schedules for the children was removed from the program design. While these supply-side issues were being dealt with, for the first 8 months of the program none of the health related conditionalities were imposed. After this period only the regular health-checks conditionality was imposed.

4.3 Model and Empirical Specification

4.3.1 Model

This paper applies three different approaches to forecasting the impact of the RPS program on different quantiles of the WAZ distribution. In an *ex ante* evaluation data is available on the untreated population. Then the unobserved counterfactual to be estimated is the outcome for the untreated group if they had been treated. In this case, the dependent variable H and

a vector of covariates X are observed for the control group. Each of these has marginal distributions F_H and F_X . The relationship between H and the covariates is assumed to be generated through a reduced form equation:

$$H = f(X, \epsilon) \quad (4.1)$$

In this paper F_H represents the marginal distribution of WAZ for the control group in the year 2002 and F_X the distribution of the covariate vector, which includes a measure of income $-F_{02}$ and school costs $-S_{02}$ for the program eligible children. The pre-treatment reduced form for the control group can then be represented as:

$$H = f(F_{02}, S_{02}, Z, \epsilon) \quad (4.2)$$

where Z includes all other covariates. Under the program the values of F_{02} and S_{02} are influenced by the policy-maker and shift with the exogenous cash transfer and, together with the Z variables, form a new covariate vector \hat{X} with a distribution $F_{\hat{X}}$. Under this new distribution of covariates ie. the distribution under treatment for the control group, the unobserved counterfactual in the *ex ante* evaluation is:

$$H^* = f(\hat{X}, \epsilon) \quad (4.3)$$

H^* is assumed to have a distribution function F_{H^*} . The objective is to estimate this unobserved distribution and compare different quantiles with the observed distribution of H , F_H .

To specify the origins of the vector \hat{X} this paper relies on the reduced form behavioural model approach proposed by Todd & Wolpin (Forthcoming). The approach relies on two key assumptions, first, that the impact of the program on the outcome variable is transmitted solely through the budget constraint and the second, that there is adequate variation in the observed policy-related variables (household income and school costs) from which to extrapolate the unobserved outcomes. The approach does not need specification of a utility function but assumes a constant set of preferences before and after the program is implemented. By using

the variation in the untreated data, untreated individuals with covariates $f(X)$ are "matched" with other untreated individuals through the function $f(\hat{X})$.

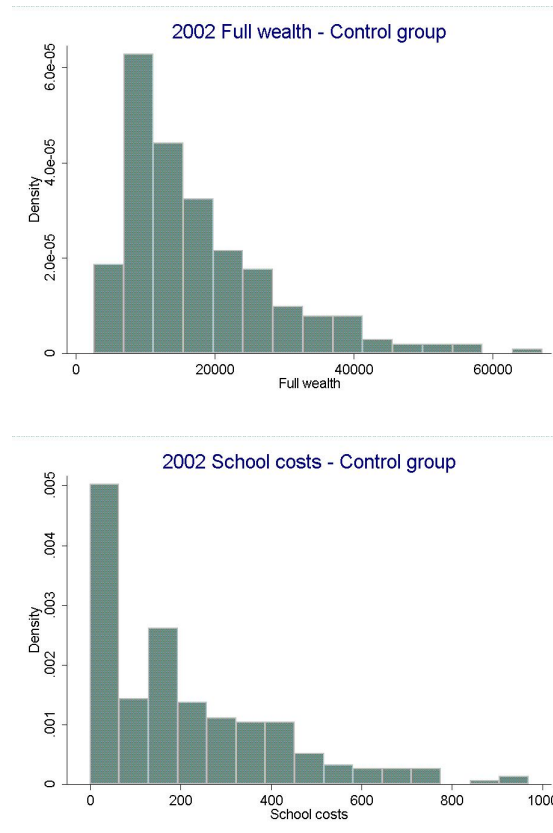


Figure 4.1: Data Variation

This exploits two sources of variation in the data to compare untreated individuals with outcomes H with other untreated individuals with outcomes H^* - the first is school costs and the second is the full income of the households at the baseline. Primary education is free in Nicaragua and most families face no fees, the cost here includes other expenditure related to schooling which is exogenous in the sense that it is faced by all families when enrolling children irrespective of whether the tuition is free or not, these include books, transport, school meals, uniforms etc. The dataset provides as a proxy for full income, a measure of household consumption. Figure 4.1(a) shows a histogram of consumption of families, with values ranging from c2618 to c67130.39; Figure 4.1(b) shows that school costs used in the estimation range from c0 to c969.

The underlying idea in the method proposed by Todd & Wolpin (Forthcoming), and the empirical strategy applied here, is similar to other non-experimental methods such as matching

which rely on 'selection on observables' (Heckman et al. 1998). The identification assumption in the methods used to estimate the reduced form is that conditional on these covariates unobserved factors are independent of the policy variables. In this paper, both components of the program are included. The school component of the program involves adjusting the indirect/compulsory costs of schooling with the lump sum school transfer. Under the assumption of full compliance, i.e. all mothers attending the compulsory health education workshops and all children being taken to scheduled health checks, the food security transfer is added to the 'full income' of the household along with the per child transfer for school supplies. By relaying the impact of the program through the budget constraint this approach models the impact of the cash transfer component of the policy. It does not allow for the impact of the different health related conditionalities on health outcomes. Isolating the contribution of the cash transfer is in itself useful as it addresses a criticism of randomized experiments - in that they operate as a 'black box' (Deaton 2010) and it is impossible to see which components produce the change. In a program such as this, with conditionalities that are likely to influence health outcomes, it then provides an *ex ante* decomposition of the contribution of one year's cash transfer to the distribution of WAZ for the year 2002. The nature of the estimation techniques used in this paper are all built around cross-sectional approaches and hence do not account for unobserved heterogeneity that could be caused by correlated unobservables. Particularly in this paper, the use of consumption as a measure of socio-economic status could be correlated with the unobservables that jointly determine nutrition status. Hence the effects detected here do not provide a causal effect of the cash transfers but can be interpreted as a partial effect of the cash transfer to the total change in the WAZ distribution.

The reduced form under the program is then:

$$H^* = f(\hat{F}_{02}, \hat{S}_{02}, Z, \epsilon) \quad (4.4)$$

Rothe (2010) refers to the above set up as a "dependent policy scenario" with a data structure $(Y_i, X_i, \hat{X}_i)_{i=1}^n$. The policy causes changes in the marginal distribution of the covariate vector that determines the WAZ score while maintaining the same conditional distribution of WAZ given X . In their paper Todd & Wolpin (Forthcoming) focus on estimating the average impact ie. the intent to treat (ITT). This paper goes beyond their approach by looking at

changes in different points of the WAZ distribution rather than just the mean. It also uses different estimation approaches that allow recovery of the differences in quantiles before and after treatment. The parameters of interest in this paper are the “quantile policy effects”:

$$\Delta_Q(\tau) = Q_H^*(\tau) - Q_H(\tau) \quad (4.5)$$

The following section describes the different estimators used. In addition, linear quantile regression is also used to estimate the results from the *ex post* randomized experiment to provide a benchmark to assess the performance of the *ex ante* estimators.

4.3.2 Empirical framework

Estimating School Costs - Two part model

The identification of the change in WAZ due to RPS relies on two policy variables - full wealth and school costs for the households. School costs are however observed in the survey only for those enrolled in school at the time of the survey and must be estimated for those not enrolled. Estimating unobserved costs typically involves using models with two components - one that determines participation ie. enrolment in school and one that involves the determinants of the cost component that is used to extrapolate the costs for the unobserved individuals in the sample. The *two-part model* assumes that the participation decision $Pr(y > 0|x)$ is determined by a parametric binary regression model either a logit or a probit, while the second part is a linear specification of x . Various specifications for the second part have been applied. Typically to deal with skewed cost data log transformations are applied. In this paper although costs are skewed a log transformation is not applicable since the observed school costs also include zeros. To deal with these issues a common specification of the second part is the *generalized linear model* which allows different distributional specifications from the exponential family to link the random with the stochastic components of the model. This however requires *a priori* specification of both the link and variance functions, in the latter case assuming certain forms of heteroskedasticity. Incorrect specifications can lead to bias and inefficiency in the estimates. An alternative is to use an approach that does not require prior assumptions about the link and variance functions. In this paper the second-part

is estimated using a semi-parametric *extended generalized linear model* (Basu & Rathouz 2005) which allows for a flexible link function and variance parameter, both of which are estimated from the data. The approach allows for a family of link functions represented by λ :

$$g(\mu_i; \lambda) = \begin{cases} \frac{\mu_i^\lambda - 1}{\lambda} & \text{if } \lambda \neq 0 \\ \log(\mu_i) & \text{if } \lambda = 0 \end{cases}$$

where g is a function that links μ to the linear predictor $X_i^T \beta$ and $\mu_i = \mu(X)$. The variance is characterized by family of functions represented by a power variance which allows for common distributions such as Poisson, Gaussian and Gamma:

$$h(\mu_i; \theta_1, \theta_2) = \theta_1 \mu_i^{\theta_2}$$

Ex post outcomes - Linear quantile regression

Results from the randomized social experiment can be used to provide a benchmark for comparison of results from the *ex ante* approach. Linear quantile regression goes beyond the mean and provides information of the impact of the program at different points of the distribution. Heckman, Smith & Clements (1997) highlight various parameters of interest that require more than the mean, including the proportion that benefit from the program and impacts at quantiles. Of particular interest in social programs such as CCTs is the impact at the lowest ends of the distribution. In the case of WAZ children with scores less than -2 are considered malnourished. Then it is important to see if the program had any impact on the weakest sections of the population by exploring heterogeneity. Focusing on the average impact could disguise potentially important benefits to those most likely to need it. The impact at a given quantile of the WAZ distribution is the vertical distance between the quantile functions in the treatment and control groups. Figure 4.2 shows the quantile plots of WAZ for the two groups in the year 2002. The graph shows that the distribution of WAZ in the treatment group is marginally higher than in the control group in the lower quantiles but they converge at the upper quantiles of the distribution. The vertical distance between the two plots at each quantile is the quantile treatment effect.

Empirically, the required impact at a particular quantile (τ) is the coefficient on the dummy

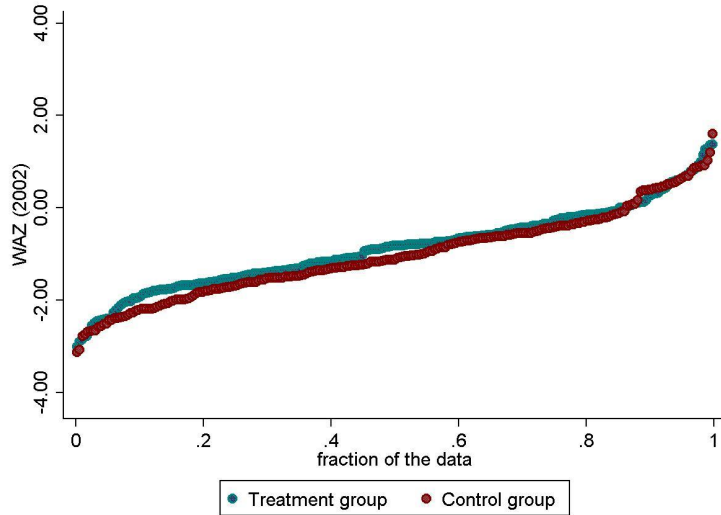


Figure 4.2: 2002 Quantile functions

variable representing participation in the program Koenker & Bassett (1978) :

$$Q_{\tau}(h|T) = \alpha(\tau) + \beta(\tau)T \quad (4.6)$$

The above equation represents the change in the distribution of WAZ in 2002 due to participation in RPS. Interpreting these results as the treatment effect for individuals requires a further assumption of "rank preservation". This requires that the rank of the potential outcome for a specific individual remains the same with and without the treatment. This however is a strong assumption and can only be tested in terms of the observables. Since this assumption does not extend to the *ex ante* methods used here, the interpretation used is a general change in the quantiles of the WAZ distribution in the treatment and control groups and no inference is made about individuals at specific quantiles of the distribution. From a policy perspective knowing the change in distribution of WAZ due to RPS is in itself informative. An upward shift would indicate an overall improvement in the WAZ score.

Quantile regression based approach

The first of the two semiparametric approaches applied in this paper is based on linear quantile regression. Proposed by Melly (2005) it involves a two step procedure to estimate the

unobserved \hat{F}_h . In the first step the conditional distribution of WAZ, given X , is estimated using linear quantile regression:

$$F_{h|x}^{-1}(\tau|x) = x'\beta(\tau) \quad (4.7)$$

where $F_{h|x}^{-1}(\tau|x)$ is the τ th quantile of WAZ conditional on the covariates. The semi-parametric nature of this estimator comes from maintaining the assumption of a linear functional form for the conditional distribution but relaxing the requirement of assuming a specific distribution. The second step involves integrating the estimated conditional distribution over the vector of covariates under the policy ie \hat{X} to recover the unconditional distribution of WAZ after treatment:

$$F_H^*(h) = \int_X F_H(h|x)dF_{\hat{X}}(x) \quad (4.8)$$

where from equation (4.7):

$$F_H(h|x) = \int_0^1 1[F_{h|x}^{-1}(\tau|x) \leq h]d\tau \quad (4.9)$$

Then for particular quantiles of WAZ the sample equivalent of the ‘‘quantile policy effect’’ is estimated by replacing $F_{h|x}^{-1}(\tau|x)$ by its consistent estimate $x'\beta(\tau)$ in equation (4.8) and averaging over all the covariates in \hat{X} . Thus the difference between the observed and the estimated counterfactual is explained by changes in characteristics. This estimator is similar to the one proposed by Machado & Mata (2005). Both estimators use linear quantile regression to estimate the conditional distribution in the first step. The Machado & Mata (2005) approach differs in the second step by using a process of random sampling from the covariate vector under the policy \hat{X} and weighting (combining) these with the vector of coefficients from the first step. Both methods have been shown to give similar results in large samples (Albrecht et al. 2009). In this paper the Melly (2005) approach is used as the sample is quite small and the repeated random sampling from different quantiles for the Machado & Mata (2005) does not work well.

Distributional Regression

The second approach used in this paper relies on a method applied initially by Han & Hausman (1990) and Foresi & Peracchi (1995) to allow for flexible estimation of conditional distribution functions. In this approach families of binary response models of varying 'cut-offs' are used to estimate a binary response, modeling the conditional distribution function separately at a set of thresholds. Chernozhukov et al. (2009) replace the linear quantile regression of Melly (2005) by applying this idea to the first step of the estimator to estimate the conditional distributional function of the outcome pre-policy:

$$F_H(h|x) = \Lambda(m(h, x)) \quad (4.10)$$

where a separate binary response model is estimated for each threshold h , Λ is a known link function such as a probit or logit and $m(h, x) = (x'\beta(h))$. Each $\beta(h)$ is estimated by maximum likelihood using a set of indicator variables $1[H \leq h] : h \in H$. The second step of this estimator follows the same process as Melly (2005) and integrates over the new covariate vector to recover the marginal distribution:

$$F_H^*(h) = \int_X \widehat{F}_H(h|x) dF_{\widehat{X}}(x) \quad (4.11)$$

Similar to the linear quantile regression approach, the Chernozhukov et al. (2009) method also provides the marginal quantile function which can be used to estimate the quantile policy effect in equation (4.5):

$$Q_{\widehat{H}}(\tau) = \inf[h : F_H^*(h) \geq \tau] = Q_H^*(\tau) \quad (4.12)$$

Nonparametric Approach

The two approaches described above are both different forms of semiparametric estimators, while the first makes no assumptions about a distribution, it still assumes a linear functional

form for the conditional quantiles. The second makes clear assumptions about a particular link function but is more flexible in that it allows for a series of binary response models that approximate the unknown distribution by a step function. The last approach applied in this paper is a nonparametric estimator (Rothe 2010) which makes no assumptions about either the distribution or functional form in estimating equation (4.4). Like the others, this estimator is also a two step process. The conditional distribution function is first estimated using kernel regression:

$$\hat{F}_{h|x}(h, x) = \frac{\hat{g}_{HX}(h, z)}{\hat{f}_X(x)} \quad (4.13)$$

where

$$\hat{g}_{HX}(h, x) = \frac{1}{n} \sum_j \{I(H_j \leq h)\} K_{x,h}(X_j - x)$$

$$\hat{f}_x(x) = \frac{1}{n} \sum_j K_{x,h}(X_j - x)$$

In the above equations K is a kernel function, h represents the bandwidth. The above specification is similar to the Nadarya-Watson estimator for kernel regression commonly available in statistical software. The second step estimates the marginal distribution using sample counterparts:

$$\hat{F}_H^*(h) = \frac{1}{n} \sum_{i=1}^n \hat{F}_{h|x}(h, \hat{X}_i) \quad (4.14)$$

As this is a nonparametric estimator's extrapolation is valid only in regions of support in the data: the estimated policy effects are only valid to regions of \hat{X} that are in the support of X . Rothe (2010) shows that despite the estimator being nonparametric it is not affected by the “curse of dimensionality” problem that frequently limits the application of such estimators. The estimates of the parameters of interest are shown to converge at the usual parametric rate of \sqrt{n} irrespective of the dimension of the covariate vector X .

4.4 Data and Variables

The dataset used in this paper was collected as a panel survey for the randomized experimental evaluation. Three rounds of data were collected - baseline (2000) and two years of follow-up (2001, 2002). In total 1581 families were included in the experiment across both treatment and control groups. In addition to the main survey, in the years 2000 and 2002 an additional module was included that captured information on the health of children aged under 5. The outcome used in this paper - weight-for-age (WAZ) Z scores was collected in this module.

The analysis in this paper uses data from the 2002 follow-up survey and pre-program census survey. The analysis includes individuals from only the randomized out control group. In total the estimation sample consists of 358 households with children under the age of 5 who are eligible for the food security and health component of the program and have been in the program for at least 2 years (i.e the length of the program); of these 236 families also have children between the ages of 7-13 who have not completed grade 4 of primary school and hence were eligible for the school transfers as well. While the total number of children in the control sample in 2002 is 530, only 239 of these children were in the baseline i.e, have been in the program for the 2 years of the experiment. The analysis is restricted to this subgroup.

The surveys gathered information on individual, household and community level variables. At the individual level detailed information was collected on education and school enrolment, direct and indirect costs of all school related expenditures, illnesses in the previous six months and health care expenditure. For all children under 5 detailed information was gathered on immunization, health checks, weight measurements. In the additional module each child was weighed and heights recorded by the surveyor and entered. The survey however does not include detailed information on earnings and income, instead detailed information on all household consumption.

The unobserved distribution of WAZ under treatment, $F_H^*(h)$, is estimated using the reduced form specification in equation (4.4). The specification includes the two key policy related

variables - consumption and school costs of the household. It also includes other variables referred to in the previous section as the vector Z , that are likely to influence the health of the child such as individual characteristics - mother's age, child's age and gender. The literature often cites "sibling effects" as a determinant of child health, where greater the number of siblings below a certain age, smaller is the amount of resources available per child. To capture these effects, number of children aged under 5 in the household is included. To capture time-related costs of accessing health care two community level variables - distance to the nearest form of public transport and travel time in hours to the nearest nurse are included. Most areas in the program have limited access to formal health care with the most likely being a health post or a trained nurse. Also included in the specification is the travel time in hours to the nearest pharmacy which is expected to capture the effect of treatments not captured by visiting a health care worker. If the time costs for parents are high due to waiting times to see a health care worker then parents might choose over-the-counter medication as a first point of medical care. The measure of "full wealth" included is expected to capture effects of having access to clean water and sanitation which are largely determined by economic status and hence these two variables are not included separately in the model. Following this argument the main specification does not include a variable to capture the education of either the mother or the head of the household. Both these variables have a large number of zeros. To test for specification robustness, the models are re-estimated including education of the household head.

The school costs for all children in the family eligible for the school components are estimated using a two part model with the same specification of variables for both parts. Due to the small sample size the model is estimated jointly for boys and girls. The variables include characteristics of the household head such as age, gender, employment and years of education. Child related variables - age and gender, and household composition variables - number of children aged under 5, number of children between 7 and 13 and the number of adults. Log of household expenditure is included as a measure of economic status. In addition, as a measure of opportunity and time costs of schooling, the distance of the household to the nearest primary school is included. The variable number of children aged under 5 in the household is also an indicator of the opportunity costs which could involve caring for younger siblings in the household.

4.5 Results

4.5.1 School costs

As discussed in the earlier sections, the first step was to estimate the school costs for the children eligible for the school transfer but not enrolled in school. The results from the *two part model* are listed in Table 4.1. Column 1 provides the estimates for the *probit* participation model for school enrolment. The binary indicators of different age groups show that younger children are more likely to enrol in school. The coefficients are positive and statistically significant. As the children grow older (age 13) they are less likely to enrol, as is common in developing countries and possibly drop out to either begin work or stay home to care for younger siblings in the household as is reflected in the negative coefficient on age 13. The coefficient on gender provides more insight and shows that boys are less likely to enrol in comparison to girls. The negative and statistically significant coefficient on gender indicates that boys may be beginning employment earlier. Most of the households in the sample cultivate lands and the gender difference may also be related to the negative coefficient on employment of the household head. This variable largely reflects land cultivation for coffee in Nicaragua which is the primary occupation of most of these households. It could mean that boys may begin to work in the fields much earlier than girls. The results show that household wealth (as measured by expenditure and asset ownership) has a positive and statistically significant effect on enrolment, with households having greater wealth being more likely to enrol. Also important is the composition of the household. Families with children under the age of 5 are less likely to have older children enrolled in school. This variable is included as a measure of the opportunity costs of schooling where older children if not employed may be expected to provide care for younger siblings. A similar negative and statistically significant effect is observed for distance to school, with children being less likely to enrol if schools are far away. The coefficient on the variable- number of children of the same age (ie 7-13 years) in the household is positive. This possibly indicates a sharing of resources that enables them to participate in school. The other key determinant of enrolment in the model is the years of education of the household head, with a positive coefficient that is statistically significant; the greater the levels of education of the household head, the higher is the likelihood of children enrolling in school. The age of the household head and gender ie. if the household head is

male, have a positive influence on the likelihood of children in this age group enrolling in school.

The second column of Table 4.1 shows the results from the *generalized extended linear model* or the *extended estimating equations approach* of Basu & Rathouz (2005). As can be expected, school costs increase with age and the coefficients are statistically significant with children aged 11 having the largest school costs. While children aged 12 and 13 tend to have marginally lower school costs they are still greater than for children below 11. Wealthier families spend more on education as do families where the household head has more years of education. School costs are lower for households where there are children between the ages of 7-13 reflecting economies of scale. Children of the same age group may be likely to share limited resources available including books and supplies across classes and age groups. A similar negative effect is observed for children under 5 years and could mean that expenditure on education is constrained by large family size and could reflect the opportunity costs of schooling as observed in the participation model for enrolment. Gender, age and employment of the household head all have negative effects on school costs but are not statistically significant. Expenditure tends to be lower when women are the head of the household. The negative coefficients on the age and work of the household head could once again reflect opportunity costs of working on the land if the household head is elderly or employed in cultivation.

As discussed in the section on empirical specification, the *extended estimating equations approach* allows for a flexible link and variance function, both of which are estimated from the data. The value of $\lambda = 0.914$ is close to an identity link function while $\theta_1 = 0.693$ and $\theta_2 = 1.246$ together are close to a Gamma distribution which requires $\theta_1 > 0$ and $\theta_2 = 2$.

4.5.2 Ex post - distributional impact

The quantile effects provide information on how the impact of the cash transfer varies at different points of the WAZ distribution. Figure 4.3 plots the quantiles using post-treatment

Table 4.1: Estimates of the model for School Costs

VARIABLES	(1) Probit Enrolment	(2) EEE School Costs
age8	0.495** (0.191)	-0.004 (0.073)
age9	0.383** (0.169)	0.181** (0.072)
age10	0.581*** (0.187)	0.163* (0.058)
age11	0.720*** (0.212)	0.340*** (0.107)
age12	0.450** (0.989)	0.195** (0.091)
age13	-0.024 (0.221)	0.222* (0.118)
Gender	-.310*** (0.115)	-0.088 (0.054)
Full Wealth (log)	0.292** (0.125)	0.632*** (0.088)
Distance to school	-0.008*** (0.002)	-0.001 (0.001)
No. of adults	0.035 (0.414)	-0.001 0.024
Children under 5	-0.217*** (0.074)	-0.022 (0.045)
Children 7-13	0.127* (0.069)	-0.092** (0.036)
HHH gender	0.061 (0.275)	-0.088 (0.112)
HHH age	0.021 (0.007)	-0.001 (0.004)
HHH yrs of ed	0.168*** (0.047)	0.028 (0.023)
HHH works	-0.245 (0.293)	-0.149 (0.134)
Constant	-2.170 (1.217)	-5.73*** (0.809)
λ		0.914*** (0.194)
θ_1		0.693*** (0.051)
θ_2		1.246*** (0.123)
Observations	785	657

Robust standard errors in parentheses, clustered at the household level

*** p<0.01, ** p<0.05, * p<0.1

information for the year 2002. 90% confidence intervals are also plotted calculated by bootstrapping the standard errors. Also plotted as a horizontal dotted line on the graph for comparison is the average treatment effect from the double difference estimates. The average treatment effect on the treated is the difference between the changes in the treatment and control groups before and after the program was implemented. This parameter assumes a constant treatment effect across all quantiles, equal to the average value. The quantile effects in the graph show the vertical distance between the quantile functions in the treatment and control groups displayed in Figure 4.2. Changes in the distribution of an outcome at quantiles can only be classified as a 'quantile treatment effect' for an individual under the assumption of rank preservation. Without rank preservation no interpretation on treatment effects for individuals is possible. Hence the results in this paper are interpreted as the impact of the cash transfer on shifting the distribution of WAZ without any reference to individuals at different quantiles of the distribution.

Table 4.2: Estimated Policy Effects

	Mean	Q10	Q50	Q90	Q90-Q10	Q90-Q50	Q50-Q10
Total Change	0.145 (0.124)	0.33 (0.121)	0.30 (0.107)	-0.12 (0.216)	-0.45	-0.42	-0.03
	Quantile Policy Effect						
Distributional Regression		0.20 (0.119)	0.33 (0.133)	0.00 (0.226)	-0.20	-0.33	0.13
Quantile Regression		0.26 (0.102)	0.21 (0.099)	0.10 (0.103)	-0.31	-0.26	-0.05
Nonparametric Regression		0.15 (0.248)	0.18 (0.309)	-0.11 (0.307)	-0.26	-0.29	0.03

Overall the graph shows that WAZ is greater in the treatment group than in the control group with positive quantile effects in all but the highest quantile where a negative but statistically non-significant effect is observed. The greatest impact is seen at the lowest quantile and the difference decreases from the lowest percentile to the highest indicating that children with the worst WAZ scores ie those that are malnourished show the biggest improvement. Table 4.2 row 1 shows the changes between the lowest, middle and highest quantiles of WAZ. The negative magnitudes show that the inequality declines across the distribution of WAZ in both the lower (Q50-Q10) and upper (Q90-Q10) segments of the distribution. These findings suggest that the average impact does not necessarily reflect an accurate picture and

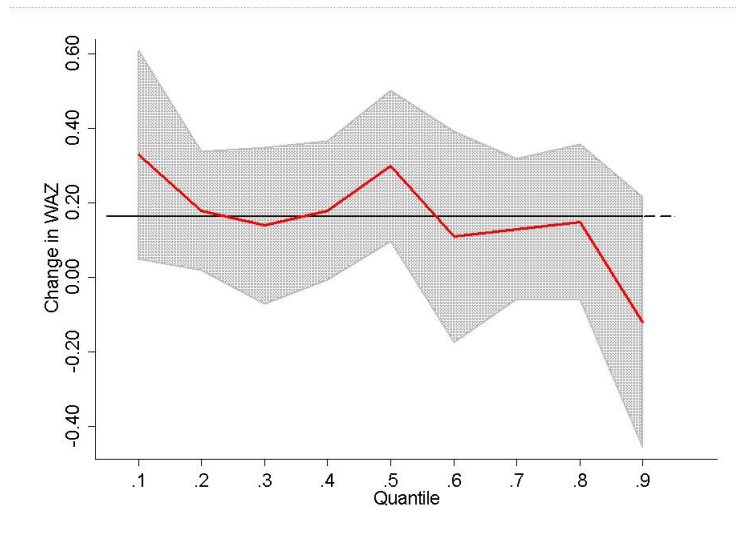
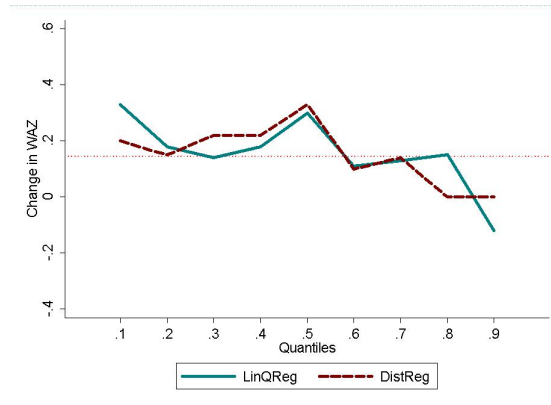


Figure 4.3: Ex-post Quantile Treatment Effects

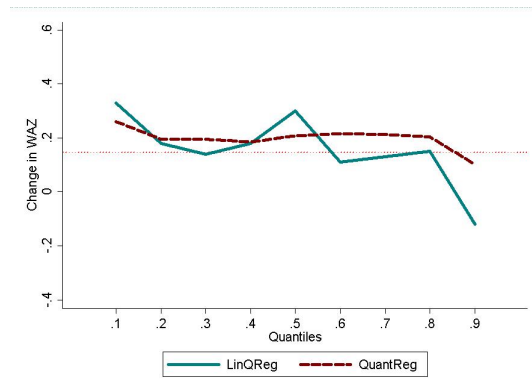
that nutrition does improve more for those who are malnourished and, as expected, little impact is seen at the higher segments of the distribution where children with the highest WAZ scores are not likely to improve any further.

4.5.3 Ex ante - distributional impact

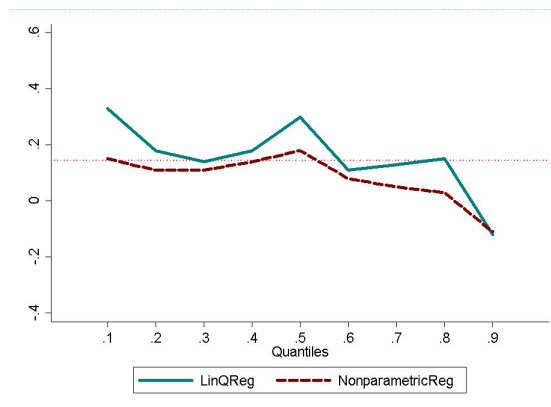
Figures 4.4(a), 4.4(b) and 4.4(c) graph the results of the *ex ante* effects of RPS on the distribution of WAZ using the semiparametric methods of Chernozhukov et al. (2009) and Melly (2005) and the nonparametric method of Rothe (2010). In all three graphs the solid line represents the *ex post* quantile effects discussed above and the horizontal dotted line represents the average effect from the double difference estimates. Overall the methods have the same direction as the *ex post* results and show the greatest impact at the lowest and middle quantiles of the WAZ distribution. Rows 2, 3 and 4 of Table 4.2 shows the results across quantiles, measured in standard deviations, for the three methods. Compared to the *ex post* statistically significant effect of 0.33 in the lowest quantile the two semiparametric methods show an improvement in child nutritional status with estimates of 0.20 and 0.26 standard deviations. In the nonparametric approach however the magnitudes of the effects are much smaller, particularly at the lowest quantile with 0.15. In the case of the 50% quantile of the WAZ distribution the distributional regression approach shows a 0.33 (statistically significant) estimate as compared to 0.30 improvement in the *ex post* case. The quantile regression



(a) Distributional Regression - Chernozhukov et al (2010)



(b) Quantile Regression - Melly (2005)



(c) Nonparametric Regression - Rothe (2010)

Figure 4.4: Ex ante Quantile Treatment Effects

approach also shows a positive impact on the lower tail of the distribution but underestimates the effect at 0.21 standard deviations while the nonparametric approach has an estimate of 0.18. All three approaches show the least impact at the highest quantiles of the WAZ distribution. While the *ex post* impact is negative and not statistically significant, the distributional regression and quantile regression approaches are also not statistically significant and show close to zero effects of RPS (0 and .10) while the nonparametric approach shows a negative and not statistically significant result. This follows the expected result where children who are not malnourished and at the upper tail of the WAZ distribution are unlikely to show large improvements from a cash transfer given to families that seem to already be investing in adequate nutrition. Table 4.2 also shows the change in the distribution across quantiles. As in the *ex post* case the quantile regression approach shows a decline in inequality in both the upper and lower segments of the distribution of WAZ. The same trend is seen for the upper quantiles (Q90-Q50) using distributional regressions and the nonparametric approach but the underprediction of WAZ at Q10 shows an increase in the Q50-Q10 segment of the distribution. This however is due to the lower predictions from these two approaches at the lowest quantile. In comparing the other quantiles (shown in Table 4.3) the distributional regression and nonparametric approaches perform well and show much closer estimates to the *ex post* case as compared to the quantile regression approach of Melly (2005) particularly at the upper tails of the distribution of WAZ. For example, Table 4.3 shows the results for the nine quantiles of the WAZ distribution. Both semiparametric approaches perform similarly for the lower quantiles of the distribution by marginally overestimating the impacts at the .30 and .40 quantiles. The two approaches however perform differently for the upper quantiles and for different specifications of the models. In the current specification the quantile regression approach over predicts the impacts across all quantiles of the upper tail of the distribution. The distributional regression approach is much closer in magnitude to the *ex post* results. For the .80 and .90 quantiles, the approach in keeping with expected results finds no impact of the RPS program on WAZ. The nonparametric method shows results quite similar to the distributional regressions and is closer in magnitude to the *ex post* results in the upper quantiles of the distribution. However, none of the estimated effects are statistically significant. The reported results for this approach are bias-corrected effects from bootstrapping the standard errors. Like the Machado & Mata (2005) approach the method is sensitive to the random draws due to the small sample size and the results are less robust than the other approaches.

Table 4.3: Estimated Policy Effects - All Quantiles

Quantiles	Q10	Q20	Q30	Q40	Q50	Q60	Q70	Q80	Q90
Total Change	0.33 (0.121)	0.18 (0.109)	0.14 (0.099)	0.18 (0.112)	0.30 (0.102)	0.11 (0.129)	0.13 (0.102)	0.15 (0.102)	-0.12 (0.237)
Quantile Policy Effect									
Distributional Regression	0.20 (0.119)	0.15 (0.112)	0.22 (0.122)	0.22 (0.161)	0.33 (0.134)	0.1 (0.095)	0.14 (0.119)	0.00 (0.232)	0.00 (0.226)
Quantile Regression	0.26 (0.102)	0.20 (0.101)	0.19 (0.102)	0.19 (0.984)	0.21 (0.099)	0.22 (0.102)	0.21 (0.108)	0.20 (0.113)	0.10 (0.118)
Nonparametric Regression	0.15 (0.248)	0.11 (0.249)	0.11 (0.276)	0.14 (0.309)	0.18 (0.300)	0.08 (0.296)	0.05 (0.263)	0.03 (0.294)	-0.11 (0.307)

As discussed earlier the methods allow recovery of the entire distribution of WAZ under treatment. This permits tests for stochastic dominance between the treated (predicted) and untreated distributions. In both the semiparametric approaches a test for stochastic dominance of a positive quantile treatment effect was carried out ie $QTE(\tau) > 0$ for all τ . The test results fail to reject the hypothesis of an improvement in the WAZ scores across all quantiles (Kolmogorov-Smirnov statistic p-values of 0.96 and 0.825 for distributional and quantile regressions respectively).

In addition to the above specification for the model a further specification including years of education of the household head is estimated. This variable includes a large number of zeros. The estimations show no change in the results of the distributional regressions but show much quantitatively larger estimates in the quantile regression approach across all quantiles, overpredicting the results across all quantiles. The nonparametric approach shows similar results to the original specification with very little change in the magnitude of the estimated effects or the standard errors

In general the three estimators provide plausibly close results to the *ex post* outcomes. The predicted results all have the same direction as the experiment. The distributional regression approach seems to perform marginally better than the quantile regression approach across all quantiles and is more stable than the nonparametric approach in small samples. In interpreting the *ex post* results it is important to recognize that they are in effect after *two years* of program intervention including the various conditionalities. If the conditionalities relating to health do have a positive effect on a short term outcome such as WAZ then the *ex ante* results should be lower than the outcomes from the experiment. This may particularly be the case at the lowest quantiles of WAZ where undernourished children are also more likely to belong to families not investing in preventive care due to resource constraints. This is reflected in the results of the lowest quantiles of WAZ. However, the *ex post* treatment effects themselves are rather small in magnitude across all quantiles. The *ex post* experimental evaluation found that 2 years of RPS decreased the percentage of underweight children ($WAZ < -2.0$) by 6.2 percentage points. Overall, the reduced form method of Todd & Wolpin (Forthcoming) does provide a close approximation of the effect of RPS on short term outcomes such as WAZ.

4.6 Conclusion

This paper combines the reduced form ex ante evaluation approach proposed by Todd & Wolpin (Forthcoming) with new empirical estimation strategies to forecast the outcomes of Nicaragua's conditional cash transfer program across the full distribution of weight-for-age Z scores for children aged under 5. The Todd & Wolpin (Forthcoming) approach relies on selection on observables and uses variation in the policy variables ie. full income and school costs of households to extrapolate the unobserved outcomes under treatment. The application in this paper however goes beyond predicting the average impact of the program and recovers the entire unobserved unconditional distribution of WAZ under the program. This facilitates forecasting the impact at different quantiles of the outcome distribution by comparing the predicted distribution with the observed pre-treatment distribution. The results are then compared with the linear quantile treatment effects from the randomized experiment.

The empirical procedure uses three different estimators - two semiparametric and one non-parametric in their specification. This first, uses linear quantile regression to estimate the conditional distribution pre-treatment and integrates this function across the distribution of the covariates under the program. The second differs by using distributional regressions to estimate the conditional distribution in the first stage - in this case a series of binary regression models are estimated at various cutoffs of the WAZ distribution. The third uses a nonparametric Kernel regression to estimate the first step and as in the other approaches averages over the distribution of the covariates under treatment. All three procedures provide estimates of the unconditional distribution of the outcome under treatment. The specifications in the empirical section are generated by the reduced form approach of Todd & Wolpin (Forthcoming).

The results show that the models provide close estimates of the effect of the cash transfers on child health across all quantiles. The distributional regression approach seems to be more stable and more precise than the quantile regression and nonparametric approach. Tests for stochastic dominance show that the program does improve WAZ across all quantiles however, the least impact is seen as expected in the highest quantiles where children are already in good health. No assumptions are made about rank preservation and hence inference is restricted

to changes in the distribution of WAZ pre and post treatment without assumptions about the impact on specific individuals at different quantiles.

Chapter 5

Conclusion

The common objective across all three papers in this thesis is to test the behavioural model reduced form approach proposed by Todd and Wolpin and in addition, to extend their idea to measure distributional impacts. In all three cases results from the *ex post* experimental evaluation from a randomized social experiment is used as a benchmark.

Chapter 2 focusses on school enrolment for children between the ages of 7-13 who had not completed grade 4 of primary school. This was the primary objective of the RPS Program in Nicaragua. The counterfactual outcome of school enrolment was estimated as a function of covariates that included household expenditure as a proxy for full wealth, school costs incurred for the eligible children, variables to capture opportunity costs of schooling. The intent-to-treat average treatment effect is estimated by matching individuals over functions of these observable characteristics. A comparison of the *ex ante* and *ex post* outcomes shows that the matching estimator predicts school enrolment quite closely for most sub-groups. The one year cash transfer increases enrolment for boys by 19 percentage points and for girls by 21 percentage points. While the analysis by sub-groups of age shows an increase in enrolment of 15 percentage points for boys and girls between the ages of 10-13 years. In all the above categories the *ex ante* approach provides estimates very close to the results from the experimental evaluation. In one sub-category of boys and girls aged 7-9 the *ex ante* approach under predicts the impact of the program. The method relies on accurately capturing all direct and indirect costs of schooling to reflect the price effect of receiving the cash transfer for the school component of the program. The underprediction could reflect not having complete information on time use and costs of schooling and can be considered

a constraint of the approach given that it relies exclusively on extrapolating from observed variation in the policy related variables.

Chapter 3 extends the analysis in chapter 2 to health outcomes for infants. Once again binary outcomes are considered - in this case accessing health checks and full coverage of vaccination. In both cases the model provides very close estimates of the experimental impact. In addition, once the model was validated with the experimental results simulations of alternate policy scenarios were carried out. Two policy formulations were tested, the first reducing the amount of the cash transfer to 75% of the original program amount while maintaining the conditionalities and providing an unconditional cash transfer equal to just the food component of the program. These policy scenarios provide insight into how households are likely to allocate resources between preventive care for infants and education of older children. The simulations find that maintaining the school conditionality and assuming full compliance of mothers attending the health education workshops, but reducing the cash transfer has little impact on preventive care and on school enrolment. Important budget implications emerge in that the experimental impact could have been achieved at a 25% lower outlay. To test the importance of the conditionality of the school transfer a model of an unconditional transfer equivalent to the “food security, health and nutrition” component of the program was estimated. The results show a large reduction in enrolment for children below age 10 and a rise in enrolment for older children. The simulation also shows a drop in the proportion of children who were fully vaccinated. In general a trade-off emerges between enrolling older children who will no longer be eligible for the school component for primary education and delaying enrolment for younger children with a strong negative income effect for vaccination of infants.

Both these chapters focus on binary outcomes and estimate the impact of the CCT program using the matching estimator proposed by Todd & Wolpin (Forthcoming). The method appears to work well and the alternative policy scenarios show interesting trade-offs between education and preventive health care.

The final chapter (4), is an extension of their idea to forecast distributional impacts of the CCT program. It combines new empirical procedures for estimating unobserved unconditional distributions with the theoretical approach of Todd and Wolpin by using the reduced

form method to motivate the empirical strategy. Limited literature exists on heterogeneous treatment effects of social programs in developing countries. All of these are restricted to *ex post* evaluations of social experiments. This chapter is unique in combining new empirical methods with the *ex ante* literature. The underlying theoretical model and motivation is the same as chapter 3 but in the empirical section three different estimators are applied, two semi-parametric and one nonparametric. Unlike the other two chapters, in this chapter data from only the control group of the *ex post* experiment is used to allow comparisons with the two year effects of the program. The outcome of interest is weight-for-age z score for children under 5, a continuous variable and a short-term measure of child nutrition. By focussing on a short-term outcome the results are reasonably close to the *ex post* quantile treatment effects from the experiment and provide estimates in the same direction. The distributional regression approach provides the best estimates of the impact followed by the nonparametric approach, which even though smaller in magnitude captures the direction of the *ex post* results better than the quantile regression based approach. However, the nonparametric approach is less stable in the case of small sample sizes. In all cases the methods show that the program had the greatest impact on the lowest quantiles of the WAZ distribution thus improving the well-being of malnourished children. Like the *ex post* results however, the magnitudes of the impact are modest.

The evidence in this thesis reveals that the Todd and Wolpin approach does work well in forecasting the impacts of cash transfer programs (conditional or unconditional). However, several assumptions required by this approach may not always be feasible. The first of these is its restriction to programs that only affect the budget constraint. Built into this is the assumption of constant preferences before and after the program. This assumption is also made in the case of microsimulations in tax-benefit models. But while they may be applicable in such programs, not all social programs will be restricted to changes in the budget constraint. For example if the subsidy received under the program directly affects utility then the reduced form equations before and after the program will not necessarily remain the same and will require specification of the utility function and depend on the assumed functional form. Also this approach is less helpful in cases where the conditionalities themselves are likely to impact outcomes. For example if the conditionality “mothers attending health education workshops” improves health outcomes (which it likely will) in the long term, the method does not allow for capturing this behavioural dynamic. The applications in this thesis have

all been restricted to short-term impacts (1 year in the first two papers) or short-term outcomes (WAZ). The underlying idea of the Todd and Wolpin approach does not allow for cumulative effects that are best explored using dynamic models. The static models proposed by them focuses on maximising within period budget constraints and may not be appropriate for long term health outcomes with cumulative effects when cash transfers are provided over several years. In their paper, they propose matching on two year's policy related variables to forecast 2 year impacts (ie. baseline and follow up). This strategy was tested in chapter 4 (results not presented) and did provide similar results to the cross-sectional approach presented in the thesis. However, it's applicability to long-term dynamic outcomes remains unanswered. A further critical assumption relates to unobserved heterogeneity. Like the *ex post* matching estimator the one proposed for the *ex ante* case also relies on "selection on observables" and, that conditional on these covariates the distribution of any unobserved heterogeneity remains the same before and after the program. While this assumption is plausible in the case of outcomes such as school enrolment, it is less likely to provide full causal effects in longer term outcomes such as health, and hence the assumption in chapter 4 identifies a partial contribution of the program on the WAZ score.

The above discussion must however be considered alongside the benefits of this approach. The methods do provide good estimates in the short-term and are useful in isolating the impact of the cash transfer component of the programs. If used as a pre-cursor to an experiment it can provide an estimate of the effect of the cash transfer component of such programs. The inability of experiments to isolate the impacts of its different components has been frequently mentioned in the literature. In addition, it provides information on potential alternate policy formulations. In all cases these are short-term effects, but they are relatively simple to achieve in comparison to more complex structural estimation approaches. In limited resource settings this approach could provide an exploratory method for new program implementers to be followed by a non-experimental *ex post* evaluation.

References

Albrecht, J., Bjorklund, A. & Vroman, S. (2009), Unionization and the evolution of the wage distribution in sweden: 1968 to 2000, IZA Discussion Papers 4246, Institute for the Study of Labor (IZA).

URL: <http://econpapers.repec.org/RePEc:iza:izadps:dp4246>

Attanasio, O., Meghir, C. & Santiago, A. (2005), Education choices in mexico: using a structural model and a randomized experiment to evaluate progressa, Open access publications from university college london, University College London.

URL: <http://ideas.repec.org/p/ner/ucllon/http-eprints.ucl.ac.uk-14750-.html>

Barham, T. & Maluccio, J. A. (2009), 'Eradicating diseases: The effect of conditional cash transfers on vaccination coverage in rural nicaragua', *Journal of Health Economics* **28**(3), 611–621.

URL: <http://ideas.repec.org/a/eee/jhecon/v28y2009i3p611-621.html>

Basu, A. & Rathouz, P. J. (2005), 'Estimating marginal and incremental effects on health outcomes using flexible link and variance function models', *Biostat* **6**(1), 93–109.

Becker, G. S. (1965), 'A theory of the allocation of time', *The Economic Journal* **75**(299), 493–517.

URL: <http://www.jstor.org/stable/2228949>

Becker, G. S. (1975), *Human Capital*, New York: National Bureau of Economic Research.

Behrman, J. R. & Wolfe, B. L. (1987), 'How does mother's schooling affect family health, nutrition, medical care usage, and household sanitation?', *Journal of Econometrics* **36**(1-2), 185–204.

URL: <http://ideas.repec.org/a/eee/econom/v36y1987i1-2p185-204.html>

Bourguignon, F. & Ferreira, F. H. (2003), Ex-ante evaluation of policy reforms using behavioral models, in 'The Impact of Economic Policies on Poverty and Income Distribution: Evaluation Techniques and Tools', The World Bank, chapter 6.

- Brooker, S., Hall, A. & Bundy, D. (1999), 'Short stature and the age of enrolment in primary school: studies in two african countries', *Social Science & Medicine* **48**(5), 675 – 682.
URL: <http://www.sciencedirect.com/science/article/B6VBF-3VKD1V8-8/2/a13fab1f033bbb6b25dc8fa3f8b6be7c>
- Case, A., Lubotsky, D. & Paxson, C. (2002), 'Economic status and health in childhood: The origins of the gradient', *The American Economic Review* **92**(5), 1308–1334.
URL: <http://www.jstor.org/stable/3083252>
- Chamberlain, G. (1986), 'Asymptotic efficiency in semi-parametric models with censoring', *Journal of Econometrics* **32**(2), 189–218.
URL: <http://ideas.repec.org/a/eee/econom/v32y1986i2p189-218.html>
- Chernozhukov, V., Fernandez-Val, I. & Melly, B. (2009), Inference on counterfactual distributions, CeMMAP working papers CWP09/09, Centre for Microdata Methods and Practice, Institute for Fiscal Studies.
URL: <http://ideas.repec.org/p/ifs/cemmap/09-09.html>
- Cosslett, S. R. (1987), 'Efficiency bounds for distribution-free estimators of the binary choice and the censored regression models', *Econometrica* **55**(3), 559–585.
- Deaton, A. (2010), 'Instruments, randomization, and learning about development', *Journal of Economic Literature* **48**(2), 424–55.
URL: <http://ideas.repec.org/a/aea/jeclit/v48y2010i2p424-55.html>
- Djebbari, H. & Smith, J. (2008), 'Heterogeneous impacts in progress', *Journal of Econometrics* **145**(1-2), 64–80.
URL: <http://ideas.repec.org/a/eee/econom/v145y2008i1-2p64-80.html>
- Edmonds, E. V. (2004), 'How well do improvements in economic status track non-monetary measures of well-being', *Dartmouth College, Hanover, NH*.
- Engle, P. L., Black, M. M., Behrman, J. R., Cabral de Mello, M., Gertler, P. J., Kapiriri, L., Martorell, R., Young, M. E. & the International Child Development Steering Group (2007), 'Strategies to avoid the loss of developmental potential in more than 200 million children in the developing world', *The Lancet* **369**(9557), 229–242.
- Ensor, T. & Cooper, S. (2004), 'Overcoming barriers to health service access: influencing the demand side', *Health Policy and Planning* **19**(2), 69–79.

- Fernald, L. C. H., Gertler, P. J. & Neufeld, L. M. (2008), 'Role of cash in conditional cash transfer programmes for child health, growth, and development: an analysis of mexicos oportuidades', *The Lancet* **371**, 828–37.
- Fernández, A. I. & Rodríguez-Poo, J. M. (1997), 'Estimation and specification testing in female labour participation models: Parametric and semiparametric methods', *Econometric Reviews* **16**(2), 229–247.
- Foresi, S. & Peracchi, F. (1995), 'The conditional distribution of excess returns: An empirical analysis', *Journal of the American Statistical Association* **90**(430), pp. 451–466.
URL: <http://www.jstor.org/stable/2291056>
- Gerfin, M. (1996), 'Parametric and semi-parametric estimation of the binary response model of labour market participation', *Journal of Applied Econometrics* **11**(3), 321–339.
- Gertler, P. (2004), 'Do conditional cash transfers improve child health? evidence from progesa's control randomized experiment', *American Economic Review* **94**(2), 336–341.
URL: <http://ideas.repec.org/a/aea/aecrev/v94y2004i2p336-341.html>
- Glassman, A., Todd, J. & Gaarder, M. (2007), Performance-based incentives for health: Conditional cash transfer programs in latin america and the caribbean, Working Papers 120, Center for Global Development.
URL: <http://ideas.repec.org/p/cgd/wpaper/120.html>
- Glewwe, P., Jacoby, H. G. & King, E. M. (2001), 'Early childhood nutrition and academic achievement: a longitudinal analysis', *Journal of Public Economics* **81**(3), 345 – 368.
URL: <http://www.sciencedirect.com/science/article/B6V76-43GH12C-1/2/07a43a0cda39f77d624dc50c0126b96a>
- Glewwe, P., Kosch, S. & Nguyen, B. (2004), Childnutrition, economic growth, and the provision of health services in vietnam, in D. Dollar, P. Glewwe & N. Agrawal, eds, 'Economic Growth, Poverty, and Household Welfare in Vietnam', World Bank, Washington D.C.
- Grantham-McGregor, S., Cheung, Y. B., Cueto, S., Glewwe, P., Richter, L. & Strupp, B. (2007), 'Developmental potential in the first 5 years for children in developing countries', *The Lancet* **369**(9555), 60 – 70.
URL: <http://www.sciencedirect.com/science/article/B6T1B-4MRGVD0-14/2/f7be2638b42c35417f70f3b4ad4348c2>

- Haddad, L., Alderman, H., Appleton, S., Song, L. & Yohannes, Y. (2003), 'Reducing child malnutrition: How far does income growth take us?', *The World Bank Economic Review* **17**(1), 107–131.
URL: <http://wber.oxfordjournals.org/content/17/1/107.abstract>
- Hadler, S. C. (2004), *Vaccination programs in developing countries*, Vaccines, 4 edn, Philadelphia, PA: Elsevier Inc, chapter 55.
- Han, A. & Hausman, J. A. (1990), 'Flexible parametric estimation of duration and competing risk models', *Journal of Applied Econometrics* **5**(1), 1–28.
URL: <http://ideas.repec.org/a/jae/japmet/v5y1990i1p1-28.html>
- Hayfield, T. & Racine, J. S. (2008), 'Nonparametric econometrics: The np package', *Journal of Statistical Software* **27**(5).
URL: <http://www.jstatsoft.org/v27/i05/>
- Heckman, J. J., Ichimura, H. & Todd, P. (1998), 'Matching as an econometric evaluation estimator', *The Review of Economic Studies* **65**(2), 261–294.
URL: <http://www.jstor.org/stable/2566973>
- Heckman, J. J., Ichimura, H. & Todd, P. E. (1997), 'Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme', *Review of Economic Studies* **64**(4), 605–54.
- Heckman, J. J., Smith, J. & Clements, N. (1997), 'Making the most out of programme evaluations and social experiments: Accounting for heterogeneity in programme impacts', *The Review of Economic Studies* **64**(4), 487–535.
URL: <http://www.jstor.org/stable/2971729>
- Horowitz, J. L. (1993), 'Semiparametric estimation of a work-trip mode choice model', *Journal of Econometrics* **58**(1-2), 49–70.
- Horowitz, J. L. (1998), *Semiparametric methods in econometrics*, New York: Springer-Verlag.
- Ichimura, H. (1993), 'Semiparametric least squares (sls) and weighted sls estimation of single-index models', *Journal of Econometrics* **58**, 71–120.
- Ichimura, H. & Taber, C. R. (2000), Direct estimation of policy impacts, NBER Technical

Working Papers 0254, National Bureau of Economic Research, Inc.

URL: <http://ideas.repec.org/p/nbr/nberte/0254.html>

IFPRI (2001a), Evaluation design for the pilot phase of the nicaraguan red de protección social., report submitted to the red de protección social, International Food Policy Research Institute.

IFPRI (2005), Nicaragua:red de protección social (rps) evaluation dataset, 2000-2002, Technical report, Washington,D.C.: International Food Policy Research Institute (IFPRI) (datasets). <http://www.ifpri.org/dataset/nicaragua>.

Klein, R. W. & Spady, R. H. (1993), 'An efficient semiparametric estimator for binary response models', *Econometrica* **61**(2), 387–421.

Koenker, R. & Basset, G. (1978), 'Regression quantiles', *Econometrica* **46**(1), 33–50.

Li, Q. & Racine, J. (2003), 'Nonparametric estimation of distributions with categorical and continuous data', *Journal of Multivariate Analysis* **86**(2), 266–292.

Machado, J. A. F. & Mata, J. (2005), 'Counterfactual decomposition of changes in wage distributions using quantile regression', *Journal of Applied Econometrics* **20**(4), 445–465.
URL: <http://ideas.repec.org/a/jae/japmet/v20y2005i4p445-465.html>

Maluccio, J. A. & Flores, R. (2005), Impact evaluation of a conditional cash transfer program: the nicaraguan red de protección social, Research reports 141, International Food Policy Research Institute (IFPRI).
URL: <http://ideas.repec.org/p/fpr/resrep/141.html>

Manning, W. G. (1998), 'The logged dependent variable, heteroscedasticity, and the retransformation problem', *Journal of Health Economics* **17**(3), 283 – 295.

Marschak, J. (1953), Econometric methods for policy and prediction, in 'Studies in Econometric Method', John Wiley, pp. 1–26.

Melly, B. (2005), 'Decomposition of differences in distribution using quantile regression', *Labour Economics* **12**(4), 577 – 590. European Association of Labour Economists 16th Annual Conference, Universidade Nova de Lisboa, Lisbon, 9th - 11th September, 2004.
URL: <http://www.sciencedirect.com/science/article/pii/S0927537105000382>

Mendez, M. A. & Adair, L. S. (1999), 'Severity and timing of stunting in the first two years

- of life affect performance on cognitive tests in late childhood', *The Journal of Nutrition* **129**(8), 1555–1562.
URL: <http://jn.nutrition.org/content/129/8/1555.abstract>
- Miller, M. A. & Hinman, A. R. (2004), *Economic analyses of vaccine policies*, Vaccines, 4 edn, Philadelphia, PA: Elsevier Inc, chapter 57.
- Moock, P. R. & Leslie, J. (1986), 'Childhood malnutrition and schooling in the terai region of nepal', *Journal of Development Economics* **20**(1), 33 – 52.
URL: <http://www.sciencedirect.com/science/article/B6VBV-46RVBX4-V/2/03829e3fffabc79e1615747cf57a4617>
- Mullahy, J. (1998), 'Much ado about two: reconsidering retransformation and the two-part model in health econometrics', *Journal of Health Economics* **17**(3), 247 – 281.
- Neyman, J. (1990), 'On the application of probability theory to agricultural experiments. essay on principles. section 9.', *Statistical Science* **5**(4), 465–472. Trans. Dorota M. Dabrowska and Terence P. Speed.
URL: <http://www.jstor.org/stable/2245382>
- O'Donnell, O., Nicolás, n. L. & Van Doorslaer, E. (2009), 'Growing richer and taller: Explaining change in the distribution of child nutritional status during vietnam's economic boom', *Journal of Development Economics* **88**(1), 45–58.
URL: <http://ideas.repec.org/a/eee/deveco/v88y2009i1p45-58.html>
- Rothe, C. (2010), 'Nonparametric estimation of distributional policy effects', *Journal of Econometrics* **155**(1), 56 – 70.
URL: <http://www.sciencedirect.com/science/article/pii/S0304407609001924>
- Rubin, D. B. (1974), 'Estimating causal effects of treatments in randomized and non-randomized studies.', *Journal of Educational Psychology* **66**, 688–701.
- Strauss, J. & Thomas, D. (2008), *Health over the Life Course*, Vol. 4 of *Handbook of Development Economics*, Elsevier, chapter 54, pp. 3375–3474.
URL: <http://ideas.repec.org/h/eee/devchp/5-54.html>
- Todd, P. E. & Wolpin, K. I. (2006), 'Assessing the impact of a school subsidy program in mexico: Using a social experiment to validate a dynamic behavioral model of child

schooling and fertility', *American Economic Review* **96**(5), 1384–1417.

URL: <http://ideas.repec.org/a/aea/aecrev/v96y2006i5p1384-1417.html>

Todd, P. E. & Wolpin, K. I. (Forthcoming), 'Ex ante evaluation of social programs', *Annales d'Economie et de Statistiques* .

White, H. (2006), Impact evaluation: The experience of the independent evaluation group of the world bank, Technical Report Report No. 1111, Washington D.C, The World Bank.

World, B. (1998), Nicaragua poverty assessment: Challenges and opportunities for poverty reduction., Technical Report Report No. 20488-NI. Washington, D.C., The World Bank.