2014

The Impact of Conditional Cash Transfer Programs Under Risk-Sharing Arrangements: Schooling and Consumption Smoothing in Rural Mexico

Eun-young Shim
University of Pennsylvania, eunshim@sas.upenn.edu

The Impact of Conditional Cash Transfer Programs Under Risk-Sharing Arrangements: Schooling and Consumption Smoothing in Rural Mexico

Abstract

This paper develops and estimates a model of informal risk sharing with limited commitment that incorporates children's school attendance choices. The model is estimated using Mexican rural villages data from the PROGRESA experiment and is used to analyze how the presence of informal risk sharing influences schooling and child labor choices, as well as the effectiveness of conditional cash transfer (CCT) programs. In particular, I compare the outcomes (schooling, child labor, and consumption) generated under the informal risk-sharing model with those that would be obtained, forcing households to make choices under autarky. I evaluate the effect of alternative program designs that were also considered in other papers which did not incorporate inter-household transfers. I find that the number of years of schooling completed at age 18 is 0.5 years lower under autarky than with risk sharing. Also, the effect of CCT on schooling outcomes and welfare of households is larger under autarky than under risk sharing, and CCT increases consumption volatility under risk sharing, especially among households with young children who are subject to the future program benefit.

Degree Type

Dissertation

Degree Name

Doctor of Philosophy (PhD)

Graduate Group

Economics

First Advisor

Kenneth I. Wolpin

Keywords

CCT, Conditional cash transfer, Informal risk sharing, limited commitment, PROGRESA

Subject Categories

Economics | Labor Economics

THE IMPACT OF CONDITIONAL CASH TRANSFER PROGRAMS UNDER RISK-SHARING ARRANGEMENTS: SCHOOLING AND CONSUMPTION SMOOTHING IN RURAL MEXICO

Eun-young Shim

A DISSERTATION

in

Economics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2014

Supervisor of Dissertation

Kenneth I. Wolpin

Walter H. and Leonore C. Annenberg Professor in the Social Sciences

Graduate Group Chairperson

Goerge J. Mailath

Walter H. Annenberg Professor in the Social Sciences and Professor of Economics

Dissertation Committee

Flavio Cunha, Assistant Professor of Economics
Hanming Fang, Professor of Economics
Dirk Krueger, Professor of Economics

THE IMPACT OF CONDITIONAL CASH TRANSFER PROGRAMS UNDER RISK-SHARING ARRANGEMENTS: SCHOOLING AND CONSUMPTION SMOOTHING IN RURAL MEXICO

DEDICATION

I dedicate my dissertation to my family. This work would not have been possible without their love and support.

ACKNOWLEDGMENT

I am greatly indebted to Kenneth Wolpin, Flavio Cunha, Hanming Fang, and Dirk Krueger for their invaluable guidance and support. I have greatly benefited from discussion with Jere Behrman, Andrew Clausen, Harold Cole, Susumu Imai, Robert Jensen, Yena Park, Petra Todd, and Shing-Yi Wang. I thank the Federal Reserve Bank of St. Louis for its hospitality and support during my visit. I also thank seminar participants at Penn empirical micro workshop and empirical micro lunch, Federal Reserve Bank of St. Louis, and SWUFE International Workshop on Applied Microeconomics in 2012 for helpful comments and suggestions. This research was supported in part by the National Science Foundation through XSEDE resources provided by the XSEDE Science Gateways program (TG-SES130022). All errors are mine.

ABSTRACT

THE IMPACT OF CONDITIONAL CASH TRANSFER PROGRAMS UNDER RISK-SHARING ARRANGEMENTS: SCHOOLING AND CONSUMPTION SMOOTHING IN RURAL MEXICO

Eun-young Shim

Kenneth I. Wolpin

This paper develops and estimates a model of informal risk sharing with limited commitment that incorporates children's school attendance choices. The model is estimated using Mexican rural villages data from the PROGRESA experiment and is used to analyze how the presence of informal risk sharing influences schooling and child labor choices, as well as the effectiveness of conditional cash transfer (CCT) programs. In particular, I compare the outcomes (schooling, child labor, and consumption) generated under the informal risk-sharing model with those that would be obtained, forcing households to make choices under autarky. I evaluate the effect of alternative program designs that were also considered in other papers which did not incorporate inter-household transfers. I find that the number of years of schooling completed at age 18 is 0.5 years lower under autarky than with risk sharing. Also, the effect of CCT on schooling outcomes and welfare of households is larger under autarky than under risk sharing, and CCT increases consumption volatility under risk sharing, especially among households with young children who are subject to the future program benefit.

Contents

List of Tables

List of Figures

1 Introduction

Over recent decades, many developing countries have adopted conditional cash transfer (CCT) programs as a new approach to reducing poverty and increasing the human capital of the next generation. These programs give cash grants to poor households conditional on children's school attendance. In 1997, the Mexican government initiated one of the first programs, which is called the Programa de Educación, Salud y Alimentación, or PROGRESA (the current program is named Oportunidades), and similar programs were adopted throughout Central and Latin America and, more recently, in South East Asia and Africa (Fiszbein and Schady, 2009). The randomized implementation of PROGRESA and its collection of extensive data offered researchers unique opportunities to evaluate the program's effectiveness and to study alternative program subsidy designs (Attanasio, Meghir, and Santiago, 2012; Behrman, Sengupta, and Todd, 2005; Schultz, 2004, Todd and Wolpin, 2006, among others).

Most of the existing studies of the effects of CCT programs on child schooling have modeled the behavior of households without regard to transfers between households and have focused on short-term program effects. An important reality of the lives of poor families in developing countries is that, in the absence of formal lending opportunities, they manage income risk through transfers among relatives and neighbors, often referred to as "informal risk sharing" (see Fafchamps, 2008, for the survey). There is evidence that public transfers that are targeted to specific groups, such as old age pension programs or conditional cash transfers, are shared by others through informal risk sharing and crowd out existing informal transfers (Albarran and Attanasio, 2003; Angelucci and Di Giorgi, 2009; Dercon and Krishnan, 2003; Jensen, 2004). The treatment effects estimated with experimental treatment and control groups, such as those obtained under the PROGRESA experiment, are robust as to whether informal risk sharing is taking place. However, studies that develop structural models to evaluate the effects of alternative program designs, and which assume that households face only household-level income constraints, may not fully capture program effects.[1] Moreover, most of the CCT literature focuses on short-term program effects, as the data are usually only available for a few years. The longer-term effects may be different from observed short-term effects, as households adjust their risk-sharing behavior in response to both the transfers and to additional changes in the level and variability of income arising from children's increased schooling.[2] At the subsistence level, managing income risk is crucial for the ability of households to meet their daily basic needs, and thus, understanding the change in informal risk sharing behavior is important in evaluating the welfare effect of the CCT programs. Also, financial constraint is an important determinant of a child schooling and labor choice, and these outcomes may respond to the change in the availability of informal transfers (Jacoby, 1994; Jacoby and Skoufias, 1997).

This paper develops and estimates a model of informal risk sharing with limited commitment that incorporates children's school attendance choices. My model builds on and extends the earlier informal village risk-sharing models introduced in Townsend (1994), on the limited commitment framework developed in Ligon, Thomas, and Worrall (2000, 2002), Laczó (2011), and on the large economy model of Krueger and Perri (2011). The model is estimated using data from the PROGRESA experiment and is used to analyze how the presence of informal risk sharing influences the effectiveness of CCT programs. In particular, I compare the treatment outcomes (schooling, consumption, and income) generated under the informal risk sharing model with those that would be obtained forcing households to make choices under autarky. I quantify the effect of informal risk

[1] The notable exceptions are Angelucci, Di Giorgi, Rangel, and Rasul (2010), and Angelucci, Di Giorgi, and Rasul (2012).

[2] Attanasio and Rios-Rull (2000) and Krueger and Perry (2006, 2011) show that lower income variability may reduce consumption risk sharing.

sharing and study the mechanisms through which risk sharing promotes schooling attainment. I evaluate the effect of alternative program designs which were also considered in other papers that did not incorporate inter household transfers (Attanasio, Meghir, and Santiago, 2012; Todd and Wolpin, 2006).

In the model, I consider a closed village economy, populated by a continuum of overlapping generations of households who make a collective decisions about (1) the size of the transfers given to or received by each household, and (2) which households send their child to school. The arrangement is based on the households' voluntary incentive to share their income with others, and hence, the arrangement is made in such a way that the future gain from staying in the arrangement relative to autarky is sufficient to compensate for the amount households have to share in the current period. Incorporating this so-called limited-commitment constraint, the arrangement endogenously responds to the expected and the actual changes in the autarky value driven by government transfers.[3]

Households begin their lifecycle as a married couple, and, with some probability, a child birth occurs. Children are eligible to attend school between ages of 6 and 18, during which time they can either choose to stay home, attend school, or work for a child labor wage. There is a utility cost of returning to school after an interruption of school attendance. Parents' income and child wage offers are exogenous and stochastic, and household income is pooled. Children aged 19 and over are assumed to work, and household income depends on child schooling attainment. Household income may also differ across households according to permanent unobserved types of parents and children which are incorporated into the model.

Households have an incentive to participate in risk sharing mainly for two reasons. First, households are risk averse and want to smooth their consumption against future

[3] Albarran and Attanasio (2003) find evidence that the limited-commitment constraint is present in these villages.

shocks. Second, schooling interruptions are costly, so parents have an incentive to keep their child in school even in periods where they receive negative income shocks.

Directly solving for the allocation for all households for all possible states is not feasible. However, a solution can be found by solving a contract between a household and an infinitely-lived, risk-neutral agent (see Atkeson and Lucas, 1992, 1995; and Krueger and Perri, 2011). The advantage of this algorithm is that it reduces the original problem to a recursive cost minimization problem of the risk-neutral agent, who is committed to deliver a certain level of lifetime utility (which is promised in the contract) to a given household without having to track the distribution of the allocations received by all households. This approach requires assuming that the village economy is stationary.

In 1997, the Mexican government launched a large-scale cash transfer program targeting the rural poor. The program, which was called PROGRESA, paid a cash grant to mothers upon the school attendance of their children, and detailed panel data were collected to evaluate the program. The data includes 25,000 households in 506 villages, and all households in selected villages were surveyed. Thus, I observe the information of all households that would participate in the risk-sharing arrangement. The data contain detailed information about household income and consumption and also about the age, completed school grade, main activity, and wage income earned by each household member. Among the 506 villages, 186 were selected as control villages, and the CCT program for them was delayed until the end of 1999. In estimation, I use the 11 largest control villages, which have more than 100 households each.[4] The model parameters are estimated by the simulated method of moments, matching the observed income, consumption, and child activity choices predicted by the model to those in the PROGRESA data.

[4] I restrict the sample to large control villages for two reasons. First, the model assumes a continuum of households, which is approximated better by large villages. Second, the model solution algorithm the economy to be stationary. However, due to introduction of the transfer program, in the treatment villages the village economy may be in a transition path to a different equilibrium than the one in the control villages.

Based on the estimated model, I find that informal risk sharing promotes more regular school attendance for children and helps households smooth consumption against income shocks. When households engage in risk sharing, school attendance and child labor choices are not responsive to parental income fluctuations, whereas in autarky, they are significantly affected by them. A temporary decrease in parental income leads to a school enrollment rate that is 12 percentage points lower under autarky than with risk sharing. Due to high estimated adjustment costs, only a small fraction of children return to school once schooling is interrupted. Under autarky, the impact of a low income shock in one year is persistent. I find that the amount of schooling completed by children by the age of 19 is 0.47 years less under autarky than with risk sharing.

I also introduce a CCT program that is similar to PROGRESA to the estimated model with and without risk sharing and find that the effect of the CCT program on schooling outcomes and household welfare is larger under autarky than under risk sharing. Compared to when there is no CCT program, the amount of schooling completed by children by the age of 19 increases by 1.12 years under autarky and by 0.85 years under risk sharing. School enrollment rate increases by 23 percentage points under autarky and by 19 percentage points under risk sharing. One of the benefits of CCT is that it mitigates the effect of negative income shocks on household consumption and school attendance. This effect, however, is negligible under risk sharing because the allocations under risk sharing were close to efficient allocations even without the CCT. Moreover, the correlation between household income and consumption under risk sharing increases by 13 percent compared to when there is no CCT. This is because beneficiary households can rely on the CCT benefit if inter-household transfer becomes unavailable, and this weakens their voluntary participation incentive to remain in the risk-sharing arrangement. This offsets some of the welfare gains from the CCT under risk sharing.

This paper relates to a large literature that evaluates the effects of the PROGRESA program, especially to Attanasio, Meghir, and Santiago (2012) and Todd and Wolpin (2006), both of which use a structural modeling approach to model schooling decisions. Both papers evaluate the effects of alternative policies that involve reallocating program benefits among households. My modeling framework is much different in that it is an equilibrium model that allows for transfers among households. I examine how policy effects differ in an environment where households make or receive transfers and jointly make schooling decisions relative to an autarky environment.

Other closely related papers are Angelucci, Di Giorgi, Rangel, and Rasul (2010) and Angelucci, Di Giorgi, and Rasul (2013). These papers find that the PROGRESA program was effective in increasing child schooling only for the households which have extended family network ties within the same village, suggesting that informal risk sharing plays an important role. This paper takes an alternative approach by developing and estimating a model that allows the amount of risk sharing within each village to be endogenously determined. This approach is more flexible when it comes to an evaluation of alternative policies by allowing risk sharing outcomes to vary depending on the nature of each policy.

Several papers structurally estimate dynamic risk-sharing models with limited commitment using village-level data. Ligon, Thomas, and Worrall (2002) find that estimated model explains the consumption dynamics in rural Indian villages better than alternative models, including autarky and a static risk-sharing model. Laczó (2011) formally applies statistical tests to different risk sharing regimes using the same Indian villages data and finds that limited commitment with heterogenous preference outperforms alternative models. She also studies the effect of a counterfactual public cash transfer policy on the risk-sharing arrangement. Their analysis is confined to consumption risk sharing and does not incorporate schooling or any other investment decision.

Ligon, Thomas, and Worrall (2000) consider a limited-commitment risk-sharing model with private savings that cannot be seized as collateral. In a numerical simulation, they show that the existence of savings helps to smooth aggregate shocks, but it may decrease welfare by making the outside option more attractive. My model does not incorporate savings (very few households in the data have savings), but I instead incorporate the schooling decision as a form of investment. Schooling differs from savings or other physical capital because schooling investments have to be made continuously during a specific interval of the lifecycle, and households receive direct utility from schooling. An implication of my model is that households which want to invest in schooling demand more risk sharing.

This paper is also related to a large literature on informal risk sharing in rural village economies that originated with Townsend (1994), who studied risk sharing among villagers in rural India. Since then, numerous papers have documented similar behaviors in many different countries (Dercon and Krishnan, 2003; Fafchamps and Lund, 2003; Kinnan, 2011; Mazzocco and Saini, 2009; and Morduch, 1995, among others). Cox and Fafchamps (2007) and Fafchamps (2008) provide extensive reviews of the literature on informal risk sharing and evidence of crowding out of public transfers found in the literature.

The rest of the paper is organized as follows. The model is presented in section 2. Section 3 provides a description of the data. Section 4 introduces estimation procedure and results. Section 5 presents the description and the analysis of counterfactual experiments, and section 6 concludes.

2 Model

In the following subsections I first set up the model environment, including the village and household demographics and the choices and risks which these households face. Then I formally define an allocation under risk sharing with limited commitment. To solve for an efficient allocation under risk sharing, I introduce an agent who mediates interaction among households. The agent's problem is presented at the end of this section.

2.1 Set-up

2.1.1 Village and Household Demographics

I consider a closed village economy, populated by a continuum of overlapping generations of households. The life of a household begins without a child. A household is endowed with some level of lifetime utility ω_0 and unobserved type $ptype$. Fertility shocks arrive once a year, and a childbirth occurs with probability π_b. Once a child is born, the fertility process ends. Children between the ages of 6 and 18 are eligible to attend school. After a child reaches age 19, he or she is no longer eligible to attend school, and instead always works and contributes to household income. Children in this stage also may permanently leave their households and stop contributing to household income with probability π_m. This event is called a "migration" in the remaining of the paper. Households in this stage (i.e., a household whose grown-up child is no longer present), whose grown-up child is no longer present, stochastically dissolve with probability π_d, after which they do not consume nor obtain any utility. Dissolved households are replaced by new households so as to keep the village demography stationary. At any given time period, each village is populated by households in different stages in their life cycle.

A period in the model equals one school semester (six months), and there are two periods within one year (fall and spring semester). All demographic transitions including birth, migration, dissolution, and the evolution of a child's age occur once a year, in the beginning of the fall period.[5]

[5] The model assumes that each household has at most one child. This is restrictive because households in rural Mexico have around 3.5 children on average. The benefit of having multiple children in the model would be that it provides more heterogeneity across households in terms of their demand for risk sharing with other households. For example, households who have multiple children who can work for wages

2.1.2 School Grade Completion and Activity Choices of Children

Children between ages 6 and 11 always attend school. Between ages 12 and 18, they either stay home, attend school, or work for a wage. These activity choices are mutually exclusive. I define $d_t^j = 1$ if alternative j is chosen in period t where $j \in \{\text{home, school, work}\}$. If alternative j is not chosen, $d_t^j = 0$. d_t denotes a vector $\{d_t^{home}, d_t^{school}, d_t^{work}\}$. School enrollment choices for a child are made twice a year, at the beginning of the fall and the spring semesters.

Although children under age 12 always attend school, they may fail their school grade and need to repeat the year. Rather than modeling that each child completes or fails a grade every year, the model determines each child's completed grade at the time the child turns 12. Define X_t^c to be the completed grade of a child at the beginning of t. The probability that $X_t^c = X$ for $X = 0, ..., 6$ is denoted by $\pi_{gr}^{12}(X, ptype)$:

$$\Pr(X_{12}^c = X) = \pi_{gr}^{12}(X, ptype)$$

and the probability sums up to 1, $\sum_{X=0}^{6} \pi_{gr}^{12}(X, ptype) = 1$.

Between ages 12 and 18, school grade completion is determined at the end of every spring semester. The probability that a grade is completed by a child is a function of school attendance choices in that year, d_{t-1}^{school} and d_t^{school}, and is denoted by $\pi_{gr}(d_t^{school}, d_{t-1}^{school})$.

have better self-insurance ability, which will decrease their voluntary participation incentive to share risk with others. Also, the maximum amount they can receive when a conditional cash transfer program (CCT) is introduced will vary depending on the number of the children who are eligible to receive the program benefit, which implies that the effect of CCT on autarky values differ across households. However, allowing multiple children in the model is computationally burdensome; therefore, in the current stage of this study, I limit my model to a single child.

2.1.3 Type Probability

The probability that a parent or a grown-up child whose completed grade is X is high-income type is given by:

$$\Pr(type = \text{high}|X) = \frac{exp\left(\mathbf{X}\gamma\right)}{1 + exp\left(\mathbf{X}\gamma\right)},$$

where

$$\mathbf{X}\gamma \equiv \gamma_o + \gamma_X X + \gamma_6 \mathbf{1}\{X \geq 6\} + \gamma_9 \mathbf{1}\{X \geq 9\} + \gamma_{12} \mathbf{1}\{X \geq 12\}.$$

2.1.4 Income Process

Households face a stochastic income process. This is a major source of risk in the model. Before a child becomes 12 years of age, his or her parents are the household's only income earners. The parental income process is given by the following:

$$log\left(y^p_{i,v,t}\right) = \beta^p_v + \mu^p_{ptype_i} + \varepsilon^p_{i,v,t} \quad , \quad \varepsilon^p_{i,v,t} \sim N\left(0, \sigma^2_{p,v}\right),$$

where β^p_v is a village fixed effect, $\mu^p_{ptype_i}$ denotes an unobserved parental component type $ptype_i$, and $\varepsilon^p_{i,v,t}$ is an idiosyncratic shock that is independent over time and across households. Let a denote the age of a child. The child's wage process between ages 12 and 18 is given by the following:

$$log\left(w_{i,v,t}(a)\right) = \beta^w_a + \varepsilon^w_{i,v,t}, \, \varepsilon^w_{i,v,t} \sim N(0, \sigma^2_w),$$

where β^w_a is an age specific mean and $\varepsilon^w_{i,v,t}$ is an idiosyncratic shock. When a child is between ages 12 and 18, household income, $y^h_{i,v,t}$, is the sum of the parental income and the child's wages:

$$y^h_{i,v,t} = y^p_{i,v,t} + w_{i,v,t}\mathbf{1}\{d^{work}_t = 1\},$$

$\varepsilon^p_{i,v,t}$ and $\varepsilon^w_{i,v,t}$ are independent of each other. As a child reaches age 19, both the parents and child jointly generate household income:

$$log\left(y^h_{i,v,t}\right) = \beta^h_v + \mu^p_{ptype_i} + \mu^c_{ctype_i} + \varepsilon^h_{i,v,t} \quad , \quad \varepsilon^h_{i,v,t} \sim N\left(0, \sigma^2_{h,v}\right),$$

where $\mu^c_{ctype_i}$ denotes an unobserved type component of a child of type $ctype_i$. As parents and grown-up child start to jointly generate household income, both the parental and the child type enter the process. If there is no child or a child is below age 12, $y^h_{i,v,t} = y^p_{i,v,t}$.

2.1.5 Preference

The period utility of household is derived from household consumption, completed school grade and activity choices of a child:

$$u_{i,t} \quad = u_\chi(\chi_{i,t}) + u_X(X^c_{i,t}) + u_{school}(d^{school}_{i,t-1}, d^{school}_{i,t}, a_{i,t}, X^c_{i,t}) + u_{work}(d^{work}_{i,t}, a_{i,t}), \tag{1}$$

where

$$u_C(\chi_{i,t}) = \frac{\chi^{1-\eta}_{i,t} - 1}{1-\eta},$$

and $\chi_{i,t}$ is a composite good of household consumption and utility derived from the child leisure given by

$$\chi_{i,t} = c_{i,t} + (\kappa_{home,a} + \xi)\mathbf{1}\{d^{home}_{i,t} = 1\},$$

with household consumption $c_{i,t}$, utility from the child's leisure, $\kappa_{home,a}$ and ξ, where $\log\xi \sim N(0, \sigma^2_\xi)$. Note that $\kappa_{home,a}$ is age-specific. The utility from leisure can be interpreted as the child's home production, and its value may change as a child gets older.

Households are risk averse over $\chi_{i,t}$, and their period utility is represented by the constant relative risk aversion utility function with risk aversion parameter η. $u_X(X^c_{i,t})$ denotes the utility from the completed school grade, $X^c_{i,t}$:

$$u_X(X^c_{i,t}) = \kappa_6 \mathbf{1}\{X^c_{i,t} \geq 6\} + \kappa_9 \mathbf{1}\{X^c_{i,t} \geq 9\} + \kappa_{12} \mathbf{1}\{X^c_{i,t} \geq 12\}.$$

$u_{school}(d_{i,t-1}, d_{i,t}, a_{i,t}, X^c_{i,t})$ denotes the utility cost of attending school:

$$\begin{aligned} u_{school}(d_{i,t-1}, d_{i,t}, a_{i,t}, X^c_{i,t}) &= \kappa_{gap1} \mathbf{1}\{d^{school}_{i,t} = 1,\, a_{i,t} - 6 - X^c_{i,t} > 0\} \\ &+ \kappa_{gap6} \mathbf{1}\{d^{school}_{i,t} = 1,\, a_{i,t} - 6 - X^c_{i,t} \geq 6\} \\ &+ \kappa_{gap9} \mathbf{1}\{d^{school}_{i,t} = 1,\, a_{i,t} - 6 - X^c_{i,t} \geq 9\} \\ &+ \kappa_{prev} \mathbf{1}\{d^{school}_{i,t-1} = 0, d^{school}_{i,t} = 1\} \end{aligned} \tag{2}$$

The top three terms on the right hand side of the equation (2) capture the utility cost of attending school when a child is behind the standard school grade for his or her age ($a_{i,t} - 6$) by at least 1 year, more than 6 years, and more than 9 years, respectively. The fourth term captures the cost of attending school when a child did not attend the school in the previous period.

The last term of (1) captures the age-specific cost of working for a wage:

$$u_{work}(d^{work}_{i,t}, a_{i,t} = a) = \kappa_{work,a} \mathbf{1}\{d^{work}_{i,t} = 1\}$$

Households discount the future at the rate of δ.

2.2 Risk Sharing Problem

In this subsection, I formally define efficient allocations under risk sharing with limited commitment and present the agent's problem in a recursive formulation. Recall that the outside option of households is to live under autarky. Therefore, to characterize the voluntary participation constraint, I first introduce the autarky problem.

2.2.1 State Space

Let us denote a vector of shocks received by household i in period t by $s_{i,t}$. $s_{i,t}$ includes stochastic events affecting demographic transition (birth, migration, and dissolution), parental income, child wage, leisure preference, grade completion. A history of $s_{i,t}$ up to period τ is denoted by s_i^τ. Define $\Omega_{i,\tau}$ to be a state that household i arrived in at the beginning of period τ, consisting of s_i^τ and demographic characteristics including the initial ex-ante lifetime utility ω_0 and the type of parents and a child, *ptype* and *ctype*:

$$\Omega_{i,\tau} = \{\omega_0, ptype_i, ctype_i, s_i^\tau, a_{i,\tau}, d_{i,\tau-1}, X_{i,\tau}^c\}.$$

Also, denote the set of all $\Omega_{i,\tau}$ that can be reached with a positive probability in period τ as Ω_τ.

2.2.2 Autarky

Under autarky, households make an optimal choice given their household budget constraint, $c_{i,t} = y_{i,t}^h$. The only choice involved under autarky is a child activity choice. The discounted present lifetime utility of household i which arrived $\Omega_{i,t}$ in period t is denoted as $U^{Aut}(\Omega_{i,t})$. For more complete presentation of the autarky problem, see Appendix A.

2.2.3 Risk Sharing with Limited Commitment

Households have an incentive to participate in risk sharing for two reasons. First, households are risk averse and want to smooth their consumption against future shocks. Second, schooling interruptions are costly, so parents have an incentive to keep their child in school even in periods where they receive negative income shocks. To manage income risk, households participate in a risk-sharing arrangement where households pool their income and make a collective decisions about (1) the size of the inter-household transfers given to or received by each household, and (2) which households send their child to school.

Households opt out to autarky unless they obtain at least as much utility as they would without any risk sharing. Also, there is no means by which to enforce this informal contract among the households. The decision whether to remain in or leave the arrangement is made by each household in the beginning of each period after $s_{i,t}$ is realized. The only exception is for the households whose child is between ages 12 and 18. For these households, elements in $s_{i,t}$ are realized sequentially. At the beginning of each period, ε^p is realized first, and then $\{\xi, \varepsilon^w\}$ follows. The decision whether to remain in the arrangement is made after the realization of ε^p but before that of $\{\xi, \varepsilon^w\}$. Thus, they remain in the arrangement if and only if the utility from staying in the arrangement is at least as much as the utility obtained under autarky in expectation over possible $\{\xi, \varepsilon^w\}$ realizations.[6] For all households, the decision to remain in the arrangement, once made, binds until the next period ε^p is realized. It is assumed that the entire arrangement breaks down if at least one household chooses to leave the arrangement, and

[6] If the participation constraint must be satisfied after $\{\xi, \varepsilon^w\}$ is realized, the dimension of the problem substantially increases, which makes the computation too costly. This is because the state-contingent utility promise has to be made for all possible combinations of $\{\xi, \varepsilon^w\}$, whereas with the current assumption, there is only one utility promise that needs to be made, which is in expectation over the all possible combinations. See Appendix for the details related to computational challenges which arise in the agent's problem.

hence, the allocation is constrained to keep every household within the arrangement in every state in every period. An allocation under the risk sharing arrangement is denoted by $(\mathbf{c},\mathbf{d}) \equiv \{\{c_{i,t}(\Omega_{i,t}), d_{i,t}(\Omega_{i,t})\}_{\Omega_{i,t}\in\Omega_t}\}_{t=0}^{\infty}$.

Define $u(\Omega_{i,\tau},\mathbf{c},\mathbf{d})$ to be the period utility of household i in period τ under allocation $(\mathbf{c},\mathbf{d})$. The constraint to keep households in the arrangement, which will hereafter be called a voluntary participation constraint (VP), are given by the following:[7]

$$u(\Omega_{i,t},\mathbf{c},\mathbf{d}) + \sum_{\tau=t+1}^{\infty} \delta^{\tau-t} E_{\Omega_{i,\tau}|\Omega_{i,t},d_{i,t}(\Omega_{i,t})}\left[u(\Omega_{i,\tau},\mathbf{c},\mathbf{d})\right] \geq U^{Aut}(\Omega_{i,t}).$$

There is no village-level aggregate saving or borrowing technology. Thus, in every period, the sum of the consumption of all households should not exceed the sum of income of the households in the village. The village resource constraint (RC) is given by,

$$\int_{\Omega_{i,t}\in\Omega_t} c_{i,t}(\Omega_{i,t})d\Phi_t(\Omega_{i,t}) \leq \int_{\Omega_{i,t}\in\Omega_t} y_{i,t}^{h}(\Omega_{i,t})d\Phi_t(\Omega_{i,t})\,\forall t,\Omega_t,$$

where $\Phi_t(\Omega_{i,t})$ is a cumulative distribution function of $\Omega_{i,t}$ over the support Ω_t in period t. Subject to (VP) and (RC), households in a village in period t jointly maximize the sum of their discounted present value of lifetime utility:

$$\max_{\mathbf{c},\mathbf{d}} \sum_{\tau=t}^{\infty} \delta^{\tau-t} E_{s^{\tau}|s^{t}} \left\{ \int_{\Omega_{i,\tau}\in\Omega_\tau} u(\Omega_{i,\tau},\mathbf{c},\mathbf{d})d\Phi_\tau(\Omega_{i,\tau}) \right\}.$$

An allocation $(\mathbf{c},\mathbf{d})$ is efficient with respect to Φ_t if the allocation solves the problem with (RC) holds at equality. As shown in Atkeson and Lucas (1992, 1995) and Krueger

[7]For the periods when a child is between age 12 and 18, (VP) is given by the following:

$$E_{s_t|\varepsilon^p}\left\{u(\Omega_{i,t},\mathbf{c},\mathbf{d}) + \sum_{\tau=t+1}^{\infty} \delta^{\tau-t} E_{\Omega_{i,\tau}|\Omega_{i,t},d_{i,t}(\Omega_{i,t})}\left[u(\Omega_{i,\tau},\mathbf{c},\mathbf{d})\right]\right\} \geq E_{s_t|\varepsilon^p}\left[U^{Aut}(\Omega_{i,t})\right].$$

and Perri (2011), solving for efficient allocations reduces to a standard dynamic programming problem. In order to use recursive techniques, however, allocations under a risk-sharing arrangement have to be restricted to stationary allocations. An allocation is stationary if $\Phi_t = \Phi$ for all t. Extending Krueger and Perri (2011), which considered an endowment economy, the solution can be found by solving a dynamic contract between a household and an infinitely-lived, risk-neutral agent as described in the next subsection.

2.2.4 The agent's Problem

The constrained efficient allocation of risk-sharing arrangement is found by solving a dynamic contract between a household and an infinitely-lived, risk-neutral agent who is responsible for allocating resources to a given household. It is assumed that the agent has an access to an outside financial market where he or she can borrow or save at a risk-free interest rate R. A household is promised a certain level of lifetime utility in the contract. The agent determines how much of the promised lifetime utility should be delivered as a form of current consumption and how much should be postponed into the next period as a utility promise. In this subsection, I set up the agent's problem in a recursive form.

In the contract a household is promised a certain level of lifetime utility, ω_t, which will be called the "promised utility" throughout the remainder of the paper. The objective of the agent is to find an allocation that minimizes the discounted sum of the cost of delivering ω_t. The agent chooses the current period utility, h, and the state-contingent promise to be made in the next period, $\omega_{t+1}(\Omega_{t+1})$, along with a child's activity choice d_t when it is relevant. Define $C(\cdot)$ to be the inverse utility function that gives the amount of consumption necessary to deliver h_t.

The agent's problem for households whose child is above age 18 is equivalent to the standard consumption risk-sharing problem which was studied in Krueger and Perri

(2011). For these households the agent's problem is given by:

$$TV(\omega_t,\Omega_t) = \min_{h_t,\{\omega_{t+1}(\Omega_{t+1})\}_{\forall s_{t+1}}} \left\{ C(h_t) + \frac{1}{R} E_{\Omega_{t+1}|\Omega_t} \left[V(\omega_{t+1}(\Omega_{t+1}),\Omega_{t+1}) \right] \right\}, \tag{3}$$

subject to

$$\omega_t = h_t + \delta E_{\Omega_{t+1}|\Omega_t} \left[\omega_{t+1}(\Omega_{t+1}) \right], \tag{4}$$

$$\omega_{t+1}(\Omega_{t+1}) \geq U^{Aut}(\Omega_{t+1}) \, \forall \Omega_{t+1}. \tag{5}$$

When a child is of age a and the child's activity choices are involved:[8]

$$V_a(\omega_t,\Omega_t) = \min_{d_t} \left\{ V_a(\omega_t,\Omega_t|d_t^{home}=1), V_a(\omega_t,\Omega_t|d_t^{school}=1), V_a(\omega_t,\Omega_t|d_t^{work}=1) \right\}$$

where

$$V_a(\omega_t,\Omega_t|d_t) = \min_{h,\{\omega_{t+1}(\Omega_{t+1})\}_{\forall s_{t+1}}} \left\{ \begin{array}{c} C(h_t,d_t) - w_t d_t^{work} \\ + \frac{1}{R} E_{\Omega_{t+1}|\Omega_t,\mathbf{d}_t} \left[V_{a+1}(\omega_{t+1}(\Omega_{t+1}),\Omega_{t+1}) \right] \end{array} \right\},$$

subject to

$$\omega_t = h_t + \delta E_{\Omega_{t+1}|\Omega_t,d_t} \left[\omega_{t+1}(\Omega_{t+1}) \right] \forall \Omega_{t+1}$$

$$\omega_{t+1}(\Omega_{t+1}) \geq E_{\Omega_{t+1}|\Omega_t,d_t} \left[U_{t+1}^{Aut}(\Omega_{t+1}) \right] \forall \Omega_{t+1}.$$

The interest rate R is endogenously determined at the level where the agent neither bor-

[8]The agent's problem presented in this subsection is simplified. For example, I did not distinguish the problems in a fall period and a spring period to simplify the presentation. The main difference between the fall and the spring problems are, when a child is present, that the child does not get older nor the school grade is updated at the end of the fall. These events only occur at the end of the spring. Other than these, the structure of the problems are similar in the fall and the spring.

rows nor saves, and thus, (RC) holds at equality.

The extension of Krueger and Perri (2011) to this setting is not trivial. It is commonly known that introducing an investment choice to a limited-commitment model makes the constraint set non-convex. This is because the investment choice affects the outside option value (under the assumption that returns from investment are still accrued under autarky), and the convex combinations of the choices of promised utility and the level of investment jointly may not be contained in the constraint set. The agent's problem considered here, however, solves the minimization problem given a discrete activity choice. Thus, given an activity choice, such a non-convexity issue does not arise, and the minimization problem with respect to the future promised utilities is well-defined.

Major challenges in computation arise due to the presence in the problem of both discrete child activity choices and continuous choice variables $\{\omega_{t+1}(\Omega_{t+1})\}_{\Omega_{t+1}}$. The presence of the dynamic discrete choices makes value functions nondifferentiable and nonconvex with respect to promised utility ω_t in the neighborhood where the optimal child activity choice changes from one to another. As the non-differentiable and non-convex value functions in period $t+1$ enter the optimization problem in period t, the first order condition is not well defined. Unlike in a standard dynamic discrete choice model where introducing many shocks convexifies the objective function, there is one continuous choice variable that needs to be solved for each possible realization of shocks.

Grid searches are robust to the nonconvexity of the objective function; however, the trade-off between the accuracy of the solution and the computation speed is large. The accuracy of the solution is important to obtain a value for R that satisfies the village resource constraint. In numerical simulations, a coarse grid of ω_{t+1} leads to failure to finding such a value for R. With a grid of ω_{t+1} that is refined enough to find an equilibrium R, the computation takes an infeasible amount of time. Because of the non-differentiability of the objective function, efficient algorithms that are based on the first order condition cannot be applied.

Fortunately, Clausen and Strub (2013) provide an envelope theorem that is applicable in this setting. According to their envelope theorem, regardless of the presence of non-differentiable parts in an objective function, optimal solutions are always found in differentiable parts of the objective function. Based on that envelope theorem, Fella (2013) proposes an algorithm that combines an endogenous grid method, which is based on the first order condition, and a grid search. By adopting this algorithm, I achieve both accuracy and speed. A detailed explanation of the computational procedure is provided in Appendix B.

3 Data

In 1997 the Mexican government launched a large-scale cash transfer program to promote school attendance of children in poor rural areas in the country. As of 1998, a total of 35,688 villages in 28 states that had a high "marginalization index" were selected for participation. Among households in each village, the program further identified the poor households that became beneficiaries of the program (For more details, see Skoufias, Davis, and Behrman, 1999). Beneficiary households whose children were enrolled in grades 3 to 9 were eligible for educational grants conditional on the attendance of their children at school for more than 85 percent of the applicable school days. The educational grant schedule is provided in Table 1.[9]

[9] The program also provided cash transfer and nutritional supplements for infants and small children that were not contingent on schooling, though the educational grant comprised more than 75 percent of the total amount of the grant (Todd and Wolpin, 2006).

Table 1: PROGRESA monthly education grant schedule in 1998

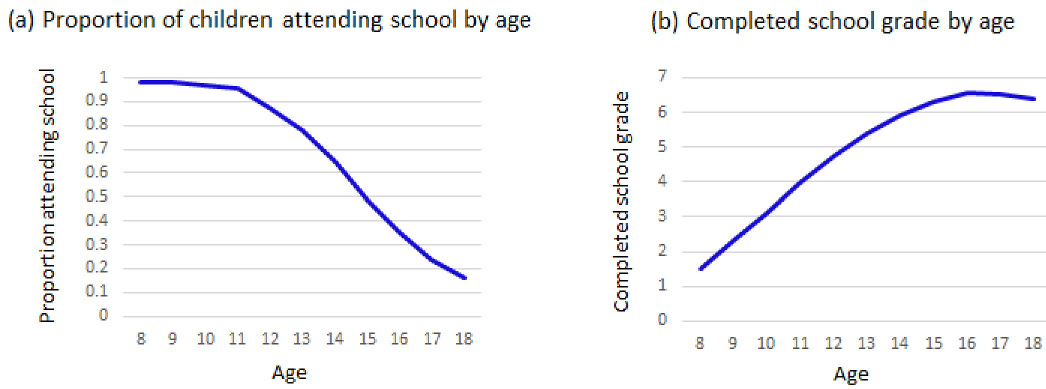
School level	Grade	Monthly grant in pesos	
		Male	Female
Primary	3	70	70
	4	80	80
	5	105	105
	6	135	135
Secondary	1	200	210
	2	210	235
	3	220	255

Source: Schultz (2004)

For the evaluation of the program, village- and household-level survey data was collected in 506 randomly selected villages. Moreover, a randomization of treatment villages and control villages was performed. The beneficiary households in 324 treatment villages started receiving the program transfers in 1998 and those in 182 control villages did at the end of 1999. The survey data was collected in both treatment and control villages before and after the implementation of the program. The baseline surveys were done in October 1997 and March 1998, and the three follow-up surveys were conducted in October 1998, May 1999, and November 1999, all before the control villages started receiving the program transfers. Every household in selected villages was surveyed; in the baseline surveys there were 24,077 households (9,221 in the control villages and 14,856 in the treatment villages). The household-level survey contains information about household members, including their age, most recently completed school grade, school attendance, employment, and information about individual members' wage income, household income from other sources, household food consumption, and household expenditure on both food and non-food items.

The school enrollment status of children at the time of the survey was collected; if a child was attending school at the time of the survey, I assume that he or she attended

Figure 1: School enrollment rate and completed school grade by age



Source: PROGRESA data October 1998, May 1999, and November 1999. The sample is restricted to all control villages.

Table 2: Proportion of children making transitions into or out of school

	All	Age≤12	Age>12
No transition into or out of school	0.72	0.80	0.43
One transition	0.18	0.12	0.38
Two transitions or more	0.11	0.08	0.20
Number of observations		16,981	

Source: De Janvry et al. (2006). PROGRESA data over 7 semesters from November 1997-November 2000.

school for the entire semester. School enrollment is almost universal before age 12 and starts to drop sharply around age 14. In control villages, the school enrollment rate between ages 16 and 18 during the data periods is around 30 percent, and the completed grade of children at age 18 is less than 7th grade (Figure 1). School attendance is not only low but also irregular. As shown in Table 2, 20 percent of children of the age of 12 and above make at least two transitions into and out of school. Table 3 shows that the most frequently cited reason for absence is “not enough money”, which suggests that financial constraint is important determinant of observed schooling outcomes in these villages.

Child labor is prevalent and a significant source of income in these villages. Child labor becomes increasingly common after the age of 12, and the fraction working for a

Table 3: Reasons for not attending school

Reason for not attending school	% who reported this reason
Not enough money to attend school	55
Did not like school	17
School is far away	6
Help was needed at home or work	6
Others	14
No response	3
Number of observations	8,527

Source: PROGRESA data October 1998, May 1999, and November 1999. The sample is restricted to all control villages and children between the ages of 12 and 18.

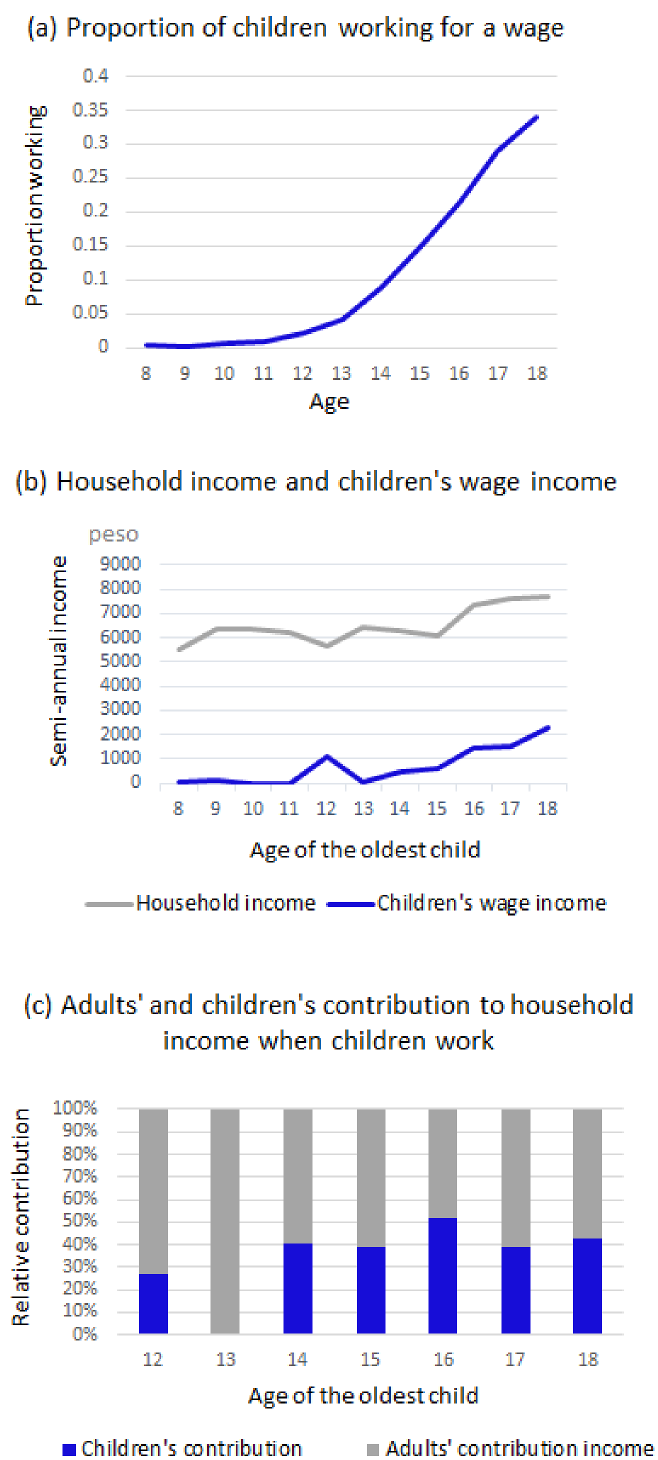
wage reaches 35 percent by the age of 18 (Figure 2(a)). The contribution of children (including zero wage) to household income increases with the age of the oldest child, reaching 30 percent of total household income at the age of 18 (Figure 2(b)). In families where a child works for a wage, the proportion of their contribution to household income is 40 percent (Figure 2(c)).

In the data, only 1.2 percent of children between the ages of 12 and 18 were both attending school and working. The assumption that the activity choices are mutually exclusive does not seem to be very restrictive.[10] Also, only 0.4 percent of the children below age 12 worked.

Table 4 shows the prevalence of shocks faced by households in PROGRESA villages. In the data, 12 percent of households reported that their head had fallen ill at least once within a given month, and almost 70 percent of them had to withdraw from work for one or more days due to the illness. Also, 16 percent of households reported that their head was unemployed at a certain point in a given month. Natural disasters such as droughts are also common. 40 percent of households reported that they had experienced a drought at

[10]When children are both attending school and working, I treat them as attending school regardless of whether they also worked.

Figure 2: Child labor in PROGRESA villages



Source: PROGRESA data October 1998, May 1999, and November 1999.
The sample is restricted to all control villages.

Table 4: Percentage of households having experienced various shocks

Household head ill at least once in a given month	12%
Withdrawn from work due to illness at least a day	68%
Withdrawn from work due to illness at least for a week	36%
Household head unemployed at least once in a given month	16%
Natural disaster in 6 months	43%
Droughts in 6 months	31%
Average percentage of households in a village reporting the droughts	40%
Number of observations	80,916

Source: PROGRESA data October 1998, May 1999, and November 1999. All villages.

least once within the past 6 months. Droughts are regarded as a community-wide shock; however, not all households in a village reported the experience of the same shock. In a given village, on average, 4 out of 10 households report that they had experienced a drought whereas 6 out of 10 households reported that they had not experienced any during the same periods. Reported damage also varied across households in each given village, suggesting that there is a room for risk sharing among households in a village even in times of natural disaster.

There is evidence that households exchange informal transfers and rely on informal sources to borrow money in PROGRESA villages. As shown in Table 5, 10 percent of households reported some form of support received from or given to other households within a single month.[11] The average amount of support was 475 pesos, which is a substantial amount considering that the average monthly household expenditure is 1,071 pesos in these villages. Also, more than 60 percent of reported credit (in frequency, not in the amount of money) was from informal sources such as relatives, neighbors, and

[11] 10 percent is arguably not large enough to support a model assumption that every household participates in risk sharing, even after considering that the survey asks about supports made during a particular month. It is possible that the phrase in the survey question, "support", may have led to under-representation of the actual extent of interhousehold transfers. Cox and Fafchamps (2007) document that in rural China, the response of households for receiving "money gifts" was five times larger than when a survey question asked about "transfers", which carried a connotation of "financial support".

Table 5: Informal financial activities in PROGRESA villages

Giving or receiving support from other households	
Percentage of households reported the activity	10%
Average amount of the support	475 pesos
Borrowing or credit	
Percentage of households reported the activity	2%
Percentage of credit from relatives/neighbors/friends	62%
Number of observations	26,972

Source: PROGRESA data October 1998, May 1999, and November 1999. Average amount of the support is reported in nominal Mexican peso.

friends.

In addition, the data on income from other sources such as sales of a household member's service, crop production, livestock sales, government program transfers other than PROGRESA (such as pensions), rent, and any interest from savings were also available. Household income variable was constructed as the sum of all members' wage income and income from other sources. Parental income is obtained by subtracting the sum of the wage income of children from household income.

The survey also contains detailed information about the quantity and the value of household food consumption, non-food expenditures, and income from various sources. The consumption variables are available in the October 1998, May 1999, and November 1999 waves, but not in the baseline surveys; thus, in my analysis, only the later three waves are used. I convert the household income and consumption to the adult-equivalent per-capita level adjusted by the economies of scale. The adult equivalent weights are based on the calculation of Di Maro (2004).[12] See Appendix C for details about how the income and consumption and variables are constructed.

[12] Di Maro (2004) assigns 1 on adult members above age 18 and 0.73 on members of age 18 and below. Then I further adjust the adult equivalent size of the households for the economies of scale by taking a square root. For example, in a household that has two adults who are age 40 and three children who are age 10, the adult equivalent size of the household adjusted for the economics of scale is $\sqrt{2 \times 1 + 3 \times 0.7} = 2.025$.

In the estimation, I use the 11 largest control villages, which have more than 100 households each. I restrict the sample to large control villages for two reasons. First, I use large villages because the model assumes a continuum of households, which is better approximated by large villages. Second, I use control villages because the model solution algorithm requires stationarity of the economy. In treatment villages, the village economy may be on a transition path to a different steady state than in the control villages, due to the introduction of the cash transfer program. There are 1,433 households in the selected sample. Also, when I map the model to the data, I only track the information about the oldest child of each household. For example, the model outcomes such as school attendance and the accepted wage of a 12-year-old-child are compared to the observed outcomes of the households whose oldest child is 12 years of age in the data.[13] Table 6 provides the summary statistics of the selected sample. The school enrollment rate and the fraction of children working for a wage by age is based on the oldest child. When constructing moments in the estimation, income is deflated so that the aggregate income level is the same as the aggregate consumption level.[14]

[13]In PROGRESA villages, the birth order does not seem to affect the school attendance decision much as the coefficient of the birth order variable in the probit regression of the school enrollment was close to zero and not significant at conventional level. The sample used in the analysis was comprised of all children between ages 6 and 18 in both the treatment and control villages (about 48,000 children in total). Other controls were age, gender, parental education, completed grade, and village-fixed effect. The most important determinant was the age of a child.

[14]The aggregate consumption-to-income ratios in each village are 0.58, 0.68, 0.96, 0.56, 0.66, 0.67, 0.77, 0.73, 1.25, 0.95, and 0.96. Except for one village, the surplus is on the income side, and in the majority of villages, the surplus is larger than 20 percent. What accounts for the surplus is not clear. Ligon, Thomas, and Worrall (2002) had the same issue in Indian ICRISAT villages data, and they suggest mismeasurement or village aggregate saving as possible sources of such discrepancies. In the estimation, I deflate income in each village so that the village aggregate consumption and income are the same.

Table 6: Summary statistics

	mean (s.d.)
Completed school grade	
mothers	3.18 (2.69)
fathers	3.81 (2.56)
children at age 18	7.42 (2.37)
School enrollment rate	
$6 \leq \text{age} \leq 11$	.98 (.13)
$12 \leq \text{age} \leq 15$	.77 (.42)
$16 \leq \text{age} \leq 18$	.36 (.48)
Fraction of children working for wage	
$12 \leq \text{age} \leq 15$	.06 (.24)
$16 \leq \text{age} \leq 18$	.31 (.46)
Children's wage income	1,179.24 (1,232.13)
Household income	2,478.75 (2,978.39)
Number of observations	4,152

Source: PROGRESA data October 1998, May 1999, and November 1999. The sample is restricted to the eleven largest control villages. Children's wage income and household income variables are in per adult-equivalent level and in 1998 Mexican pesos.

4 Estimation

I use the method of simulated moments to estimate the model parameters.[15] This estimation method finds a set of parameter values that minimizes the weighted distance between the aggregate moments constructed with the data and that of the simulated outcomes of the model. The weights are given by the inverse of estimated variances of the data moments following Lee and Wolpin (2010).

The simulated moments are generated from the computed stationary distribution for a given set of parameters. I use the observed completed school grade of mothers and the estimated type-probability process to determine the type of parents of each household. Given the parental type, the age and completed school grade of each child is drawn from the stationary distribution. I assume that the state of the economy when the parents' schooling investment was made is different from the current steady state and that the parental schooling distribution and the parental type distribution are exogenous to the model. I also assume that the new households entering the economy have the same schooling distribution as the mothers in the sample. Thus, the parental schooling distribution and the resulting parental type distribution are constant over time. Given the parental

[15] Parameters in the model are identified by the combination of functional forms, distributional assumptions, and exclusion restrictions. μ^p in the parental and household income processes for high-income type are normalized to be zero. Although the parameters in the parental and household income processes are identified outside the model, estimating the income process jointly with other model parameters makes the estimation more efficient. This is because the parameters in the income process determine the return from schooling, and thus they affect other parameters that govern child activity choices. I fix the subjective discount factor δ at a conventional level of 0.95 because with a constant relative risk aversion (CRRA) utility function it is not obvious how to separately identify the risk aversion and time discounting parameters. Grade-completion probability for grades 7-12, $\pi_{gr}(d_t^{school}, d_{t-1}^{school})$ are not estimated but imputed based on the fraction of children completing one school grade given their observed enrollment status. The computed grade-completion probability for grades 7-12 is given by $\pi_{gr}(1,1) = 0.8$, $\pi_{gr}(1,0) = \pi_{gr}(0,1) = 0.25$, and $\pi_{gr}(0,0) = 0.1$. Grade-completion probability for children below age 12 (or between grades 1-6), $\pi_{gr}^{12}(X,k)$, for each (X,k) are estimated where X is a school grade in which a child is enrolled and k is a parental type. Although $\pi_{gr}(d_t^{school}, d_{t-1}^{school})$ is independent of parental type, the proportion of children who complete a school grade each year can differ across different parental types because d_t^{school} is a choice and is affected by parental type. Demographic transition probabilities π_b, π_m, and π_d are also imputed based on the observed proportion of households which make a given demographic transition each year. The imputed values are: $\pi_b = 0.17$, $\pi_m = 0.036$, and $\pi_d = 0.12$.

type, the age and the completed school grade of a child, and a set of shocks, an efficient allocation is computed by solving the agent's problem. Then, the simulated moments are constructed as an average over the outcomes which are associated with many different sets of shocks.

Target moments used in the estimation are listed below. A "grown-up child" refers to a child whose age is above 18.

1. Income process

   (a) Mean parental income by village, excluding households with a grown-up child

   (b) Mean household income by village (households with a grown-up child only)

   (c) Standard deviation of mean-differenced log parental income by village, excluding households with a grown-up child

   (d) Standard deviation of mean-differenced log household income by village (households with a grown-up child only)

   (e) Mean parental income by the completed grade of mothers, excluding households with a grown-up child

   (f) Mean parental income by the completed grade of mothers (households with a grown-up child only)

   (g) Mean household income by the completed grade of children (households with a grown-up child only)

   (h) Standard deviation of log parental income by the completed grade of mothers, excluding households with a grown-up child only

   (i) Standard deviation of log household income by the completed grade of children (households with a grown-up child only).

2. Child wage process

    (a) Mean accepted wage by age, from ages 12 to 18

    (b) Mean accepted wage by parental income quartile

    (c) Mean accepted wage by the completed grade of mothers.

3. Child activity choice

    (a) Proportion of children attending school and working by age, from age 12 to age 18

    (b) Proportion of children staying home, attending school, and working between ages 12 and 16 by parental income quartile

    (c) Proportion of children attending in each of the 5th, 6th, 8th, and 9th grades

    (d) Proportion of children attending school in the spring but who were not enrolled in the previous fall

    (e) Proportion of children attending school in the fall but were not enrolled in the previous spring

    (f) Proportion of children attending school by years behind the standard grade level at each age

    (g) Mean parental income of households with a child attending middle school (grades 7 to 9)

    (h) Mean parental income of households whose child is eligible to attend middle school but stays home

    (i) Mean parental income of households whose child attends high school (grades 10 to 12)

(j) Mean parental income of households whose child is eligible for high school but stays home.

4. School grade completion[16]

   (a) Proportion of 12-year-old children who completed no grade, 1st grade, 2nd grade, ..., 6th grade by the completed school grade of their mothers.

5. Consumption

   (a) Mean consumption by the completed grade of mothers

   (b) Mean consumption by the completed grade of child (households with a grown-up child only)

   (c) Mean consumption by parental income quartile

   (d) Correlation of mean-differenced household income and consumption over time.

4.1 Parameter estimates

The estimated parameters are shown in Table 7-Table 10.[17] I discuss the parameter estimates of particular interests in the following subsections.

Table 7: Parameter estimates: income process

Village	Village fixed effect		Income shock variance	
	Parental income (β_v^p)	Household income (β_v^h)	Parental income (σ_v^p)	Household income (σ_v^h)
1	8.12 (.0063)	8.84 (.0210)	.88 (.0002)	.58 (.0069)
2	7.49 (.0063)	8.79 (.0370)	.87 (.0002)	.71 (.0358)
3	7.52 (.0061)	8.08 (.0331)	.90 (.0002)	.84 (.0310)
4	8.23 (.0063)	8.87 (.0325)	.57 (.0001)	.69 (.0234)
5	7.87 (.0062)	8.74 (.0274)	.55 (.0002)	.62 (.0206)
6	8.11 (.0063)	8.40 (.0390)	.87 (.0002)	.86 (.0368)
7	7.71 (.0081)	8.27 (.0456)	.70 (.0002)	.92 (.0337)
8	7.61 (.0066)	8.05 (.0288)	.79 (.0001)	.78 (.0251)
9	7.72 (.0049)	8.50 (.0391)	.71 (.0002)	.75 (.0436)
10	7.57 (.0054)	8.14 (.0308)	.89 (.0002)	1.01 (.0291)
11	7.80 (.0057)	8.62 (.0113)	.55 (.0002)	.84 (.0000)

Notes: Standard errors are in parentheses.

4.1.1 Income, wage, and type process parameters

Table 7-Table 9 provide parameter estimates for parental income, child wage, household income, and type process parameters. As shown in Table 7, the mean and the variance of parental and household income vary significantly across villages. In all villages, household income with a grown-up child is larger than that of other households. Table 8 provides parameter estimates for the child wage process. Mean wage offers for children substantially increase with the child's age, implying a higher opportunity cost of attending school for older children. Table 9 provides type process parameter estimates. The

[16]The current estimation on which the results reported in the paper are based does not incorporate the moments related to school grade completion. School grade completion probabilities $\pi_{gr}^{12}(X, ptype)$ for $X = 0, \ldots, 6$ and $ptype = low, high$ are calibrated. The results of the full estimation are forthcoming. The calibrated values are given by the following. The numbers may not sum up to 1 because they are rounded:

ptype\X	0	1	2	3	4	5	6
low	0.03	0.01	0.05	0.12	0.27	0.28	0.13
high	0.08	0.01	0.01	0.05	0.08	0.31	0.47

[17]Some of the standard errors are missing. The estimation results will be updated soon.

Table 8: Parameter estimates: child wage process

Age-specific constant (β_a^w)	
age 12, 13	2.00 (.0793)
14	4.00 (.2227)
15	4.00 (.2169)
16	4.20 (.2226)
17	5.00 (.3514)
18	5.03 (.1132)
Wage shock variance (σ_w^2)	1.62 (.0062)

Notes: Standard errors are in parentheses.

estimated type component in income process, μ^p, is -0.9 for the low-income type, translating to a mean income that is 41 percent of high-income-type ones.[18] This is compatible with the observed difference in mean incomes of poor (CCT beneficiaries) and non-poor households in the data.[19] The estimated parameters in the type probability process predict that the probability of being a high-income-type child when no grade is completed is 0.25 and that completing the 6th, 9th, and 12th grades increases this probability to 0.48, 0.56, and 0.66 respectively. Combined with the estimated μ^p, the estimated type probability process predicts that, compared to those who did not finish any school grade, the expected mean household income is 24, 33, and 43 percent higher when a child completes 6th, 9th, and 12th grades.[20] The type components of parents and children enter multiplicatively in

[18] This calculation is done by:

$$\frac{\text{Realized income of low-income type}}{\text{Realized income of high-income type}} = \frac{e^{\beta_v+\mu^p+\varepsilon^p}}{e^{\beta_v+\varepsilon^p}} == exp(-0.9) \simeq 0.41$$

[19] In the sample used in the estimation, the observed mean of per adult-equivalent household income of the beneficiary households for 6 months is 1835.68 pesos and that of the non-beneficiary households is 3022.26 pesos.

[20] The relative size of income is obtained by the following calculations:

(A) Mean income of those who did not complete any school grade = $0.75e^{\beta_v-\mu^p+0.5\sigma_v^p} + 0.25e^{\beta_v+0.5\sigma_v^p}$

(B) Mean income of those who completed 6th grade =$0.52e^{\beta_v-\mu^p+0.5\sigma_v^p} + 0.48e^{\beta_v+0.5\sigma_v^p}$

The relative size of mean income when completing 6th grade=(B-A)/A=0.24. Similar calculations were done for 9th and 12th grades.

Table 9: Parameter estimates: type process

Unobserved type component in income process	
Constant term for low-income type (μ)	-.90 (.0058)
Type probability process (probability of being high-income type)	
When completed grade is observed	
Constant (γ_0)	-1.00 (.0021)
Coefficient of completed school grade	
For grade between 1 and 6 ($\gamma_{X,primary}$)	.10 (.0143)
For grade above 6 ($\gamma_{X,secondary}$)	.08 (.0183)
Coefficient of completing primary school (γ_6)	.30 (.0595)
Coefficient of completing middle school (γ_9)	.12 (.0132)
Coefficient of completing high school (γ_{12})	.17 (.0123)
When completed grade is not observed	
Constant (γ_{X_0})	.80 (.0130)

Notes: Standard errors are in parentheses.

the household income process, and households with high-income type parents have larger return from schooling.

4.1.2 Preference Parameters

Preference parameter estimates are provided in Table 10. The estimated risk aversion coefficient is 0.63. This is within the range of the values which were found in other studies that structurally estimate limited-commitment risk-sharing models using village-level data in developing countries.[21]

The estimated parameters for utilities from completing primary, middle, and high school are 0.30, 0.50, and 0.75 respectively. To provide an idea of how large these num-

[21]Other papers that structurally estimate a limited-commitment risk-sharing model by using data from rural villages have found larger risk aversion coefficient estimates. For example, the estimated risk aversion coefficients of the CRRA utility function in Ligon, Thomas, and Worrall (2002) are between 0.84 and 1.61. These values are associated with the estimated values of time discount factors which range between 0.7 and 0.95, where a large discount factor is associated with a low risk aversion coefficient. In Morten (2013) risk aversion coefficient is fixed at 1.6 and the estimated values of discount factor lie between 0.43 and 0.66.

Table 10: Parameter estimates: preference

Risk aversion coefficient (η)	.63 (.0015)
Utility cost of attending to school	
At least one years behind of standard grade level at certain age (κ_{gap1})	-.50 (.0016)
6+ years behind of standard grade level (κ_{gap6})	-1.20 (.0016)
9+ years behind of standard grade level (κ_{gap9})	-1.20 (.0049)
Attending school after not attending school in the previous period	
Spring semester ($\kappa_{prev,spring}$)	-1.00 (.2561)
Fall semester ($\kappa_{prev,fall}$)	-.20 (.0408)
Utility from staying home ($\kappa_{home,a}$)	
For children age 14	4.50 (.1669)
For children age between 15 and 16	4.80 (.9891)
For children age between 17 and 18	5.00 (.1974)
Idiosyncratic leisure preference shock variance (σ_ξ)	3.50 (.5514)
Utility cost of working for a wage ($\kappa_{work,a}$)	
For children age between 12 and 14	-1.40 (944.29)
For children age between 15 and 16	-1.40 (.2398)
Utility from completed schooling	
Completing primary school (κ_6)	.36 (.2804)
Completing middle school (κ_9)	.56 (.2462)
Completing high school (κ_{12})	.76 (.3743)

Notes: Standard errors are in parentheses.

bers are I convert them into consumption-equivalent terms given a reference consumption level of 2,634.46 pesos, which is the estimated mean consumption level in the economy. The utilities from completing schools are translated into 183.42, 210.57, and 259.24 additional pesos of consumption respectively, compared to the households of which the child did not finish primary school.[22]

The estimated parameters for the leisure value of a child are translated into 100.34

[22]The consumption-equivalent values were obtained by the following way. Suppose that a household has household consumption of 2,634.46 pesos (which is the average amount of consumption in the villages) and that the household has a child who completed 6th grade. Let x denote the estimated utility from completing

pesos for ages 12 and 13; 272.75 pesos for age 14, 15, and 16; and 449.68 pesos for ages 17 and 18. Because consumption and leisure values are perfectly substitutable under the current specification, the utility from leisure is relatively more important for households which have a smaller amount of household consumption.

The utility cost of attending school can be converted into consumption-equivalent terms in a similar way as it was done for the utilities from completed schooling. The utility loss from attending school when a child is behind by at least one school year, 6 or more years, or 9 or more years is translated into a decrease in household consumption by 190.28, 243.70, and 298.79 pesos respectively. The cost of returning to school when a child did not attend school in the previous semester is translated into 336.12 pesos.

4.2 Model fit

Tables 11-16 provide information on how well the estimated model fits the data.

Table 11 compares the actual and predicted mean and log-variance of parental and household incomes. The model fits the mean and variance of income values at the village level very well.

Table 13 compares the actual and predicted mean and log-variance of parental income. The model fits the variance very closely, but the mean incomes for lower education groups (those with no schooling or only some primary schooling) are overpredicted.

Table 14 shows that the model fits the completed grade by age, the proportion of children attending school and working for a wage, and the mean accepted wage by age very well. In other age groups, the model predictions are very close to the actual levels

6th grade. Then, the period utility of the household is $\frac{2634.46^{1-\eta}-1}{1-\eta}+x$. I solve for c such that:

$$\frac{(2634.46+c)^{1-\eta}-1}{1-\eta}=\frac{2634.46^{1-\eta}-1}{1-\eta}+x$$

found in the data. Although the mean wage offer increases with age, because of the selection, the accepted mean wage level does not increase with age after age 13.

The predicted correlation between household income and consumption is 0.103 and is close to the actual level of 0.084 as shown in Table 16. This suggests that the allocations in these villages are close to allocations under complete risk sharing.

Table 11: Actual and Predicted Mean Parental and Household Income

	Parental income without a grown-up child		Household income with a grown-up child	
Village	Actual	Predicted	Actual	Predicted
1	3124.25	3126.74	3837.55	3829.02
2	1586.64	1527.31	4074.70	4083.43
3	1638.21	1604.79	2041.31	2024.47
4	2658.78	2628.94	4230.92	4269.64
5	1855.24	1833.06	3470.49	3493.49
6	3071.18	3035.35	2992.50	3040.96
7	1825.64	1822.71	2820.83	2836.88
8	1760.51	1757.09	1901.35	1960.59
9	1701.70	1696.31	2849.33	2822.12
10	1707.54	1714.15	2493.06	2507.88
11	1656.71	1643.14	3509.41	3526.11

Table 12: Actual and predicted log parental and household income variance

Village	Parental income without a grown-up child		Household income with a grown-up child	
	Actual	Predicted	Actual	Predicted
1	0.72	0.72	0.45	0.46
2	0.70	0.69	0.55	0.54
3	0.74	0.74	0.65	0.65
4	0.46	0.46	0.53	0.55
5	0.44	0.44	0.48	0.49
6	0.71	0.71	0.67	0.67
7	0.57	0.57	0.71	0.72
8	0.63	0.63	0.60	0.61
9	0.59	0.59	0.58	0.57
10	0.72	0.72	0.79	0.76
11	0.42	0.42	0.65	0.67

Table 13: Actual and predicted parental income by mother's completed school grade

	Mean		s.d.	
	Actual	Predicted	Actual	Predicted
no school	1812.39	2063.89	0.8757	0.8843
some primary	1946.08	2071.67	1.0164	0.8829
primary and beyond	2233.32	2287.21	0.8559	0.8829
not reported	2213.08	2156.61	1.0666	0.8934

Table 14: Actual and predicted moments by age

	Completed school grade		Fraction attending school	
Age	Actual	Predicted	Actual	Predicted
12	4.49	4.52	0.91	0.91
13	5.47	5.22	0.91	0.91
14	5.99	5.92	0.70	0.73
15	7.06	6.48	0.53	0.63
16	7.11	6.98	0.46	0.55
17	6.87	7.44	0.29	0.35
18	7.42	7.75	0.13	0.24

	Fraction working		Mean accepted wage	
Age	Actual	Predicted	Actual	Predicted
12	0.01	0.01	359.12	406.43
13	0.02	0.01	522.83	437.66
14	0.12	0.10	1503.66	1196.84
15	0.11	0.11	923.64	1159.54
16	0.17	0.12	1322.43	1320.00
17	0.32	0.32	1528.20	1448.41
18	0.44	0.49	1204.83	1042.14

Table 15: Actual and predicted consumption, proportion attending school and staying home, mean wage by income quartile

	Mean Consumption		Fraction in school	
Income quartile	Actual	Predicted	Actual	Predicted
25%	1982.09	1852.64	0.64	0.61
50%	2142.22	2231.33	0.65	0.70
75%	2259.76	2526.20	0.73	0.73
100%	2987.52	2789.29	0.78	0.78

	Mean Wage		Fraction staying home	
Income quartile	Actual	Predicted	Actual	Predicted
25%	836.45	1004.40	0.25	0.23
50%	869.86	1171.71	0.23	0.17
75%	1402.24	1226.61	0.20	0.15
100%	2150.56	1328.46	0.15	0.12

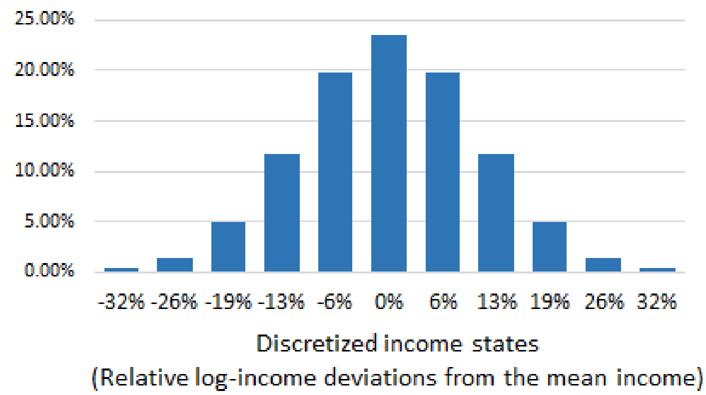
Table 16: Actual and predicted correlation between household income and consumption

Actual	Predicted
0.084	0.103

5 Counterfactual Analysis

In this section I discuss the results of the two counterfactual experiments. Throughout this section, the baseline model refers to the estimated limited-commitment risk-sharing model without any policy intervention. In the first experiment, I simulate the consumption and child activity choices under autarky given the estimated set of parameter values, and compare the outcomes to those predicted by the baseline model. The purpose of this exercise is to quantify the effectiveness of informal risk sharing in smoothing consumption and schooling. In the second experiment, I introduce CCT to both the baseline model and the autarky model. Todd and Wolpin (2006) make a prediction about the long-term effect of PROGRESA and alternative designs of the program subsidy under the autarky

Figure 3: Estimated distribution over discretized income states



assumption. Using the model developed in this paper, I can compare the predicted CCT outcomes when risk sharing interacts with CCT and when households make decisions under autarky without risk sharing. All simulated outcomes are drawn from the stationary distribution. The distribution of the completed grade of parents is not affected by CCT or the presence of risk sharing.

5.1 Risk Sharing and Autarky

Figure 4 and Table 17 provide the outcomes from the first experiment. Figure 4 shows (a) the fraction of children attending school, (b) fraction working for a wage between ages of 12 and 18, (c) average household consumption, and (d) the mean accepted wage by children between ages of 12 and 18 of households which are in different parental income states at a cross-section. The estimated cross-sectional distribution of discretized income states is given by Figure 3. On the x-axis of all four panels, the size of the relative deviations from the mean log-income which corresponds to each income state is given. The dark solid lines are outcomes obtained under risk sharing, the dark dotted lines are outcomes under autarky, and the light solid lines are from the data. If risk sharing is complete, the outcomes should not vary across different income states. As it is shown in Figure 4, in the data, the observed outcomes do not vary across different parental income states which are associated with negative or small positive shocks suggesting that households are well insured against these shocks. The outcomes under risk sharing capture this pattern very well. However, in the data, the observed schooling and consumption outcomes are positively associated with parental income states for large positive income shocks. The outcomes under risk sharing fail to capture this pattern and predict that households are well insured against large positive shocks as well. Under autarky, all outcomes vary substantially with parental income states.

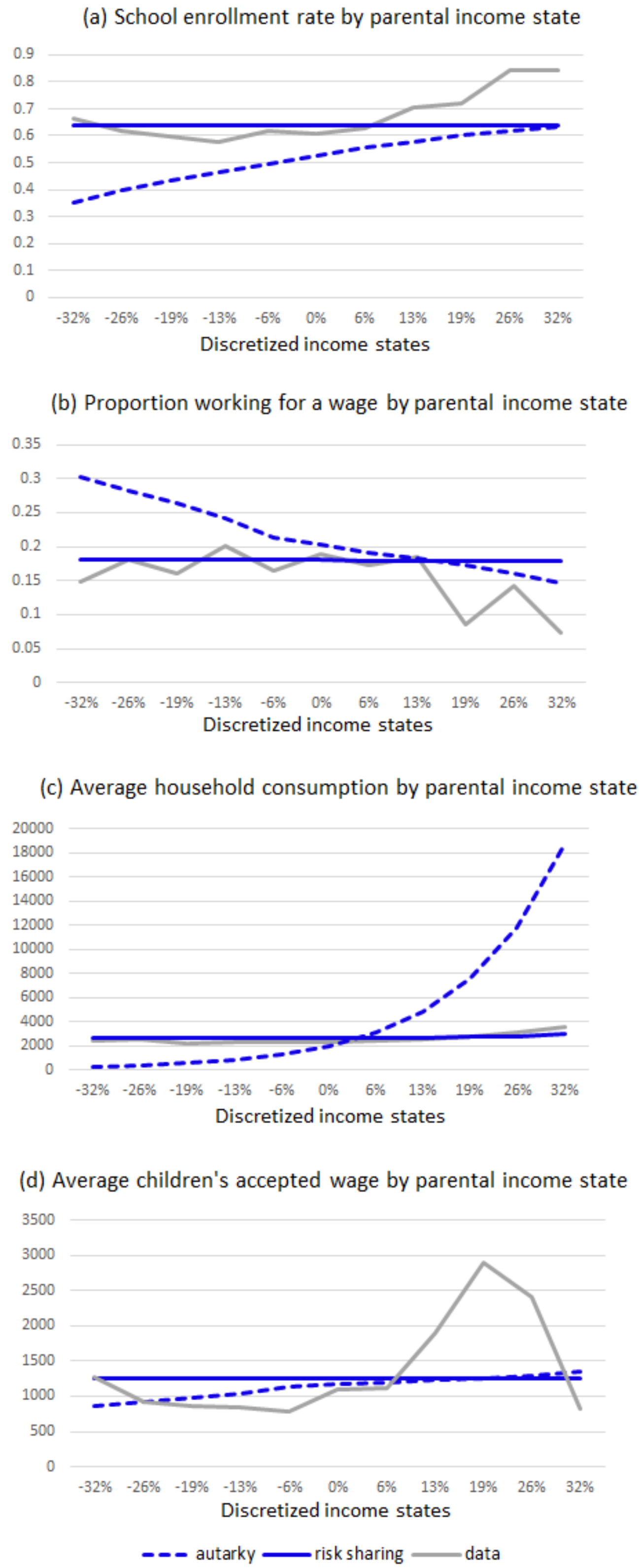
Table 17 shows that the overall fraction of children attending school is lower under autarky than under risk sharing. Under autarky, there are more children whose school attendance was interrupted at least once in the past due to a negative parental income shock. Once school attendance is interrupted, the cost of attending school becomes higher. The effect of this transitory income shock accumulates over time, which leads to a lower school attendance rate under autarky. Overall, under autarky, the fraction attending school is 11 percentage points lower than that with risk sharing and the completed school grade by age 19 is higher by almost a half year, as shown in Table 17.

Although the income earned by children younger than age 19 is smaller as fewer children work for wages under risk sharing (Table 17), a higher completed school grade leads to a larger fraction of high-income type children, which increases the size of the village

Table 17: Outcomes under risk sharing and autarky

	Risk sharing	Autarky
Completed grade at age 19	8.42	7.95
Fraction attending school, 12≤age≤18	0.64	0.52
Fraction working, 12≤age≤18	0.18	0.21
Children's (accepted) wage (in peso)	1,255.80	1,141.65
Household consumption (in peso)	2,643.06	2,627.05
Welfare in consumption equivalence (in peso)	2,034.84	1,862.54

Figure 4: Outcomes under risk sharing and autarky by cross-sectional income states



resource available for consumption. I also compare the lifetime welfare of new households under the two regimes. Welfare is converted into consumption-equivalent terms, assuming that households live for 60 years (or 120 model periods) and that the amount of consumption in every period is constant.[23] The consumption-equivalent lifetime utility under risk sharing is 2,034.84 pesos, which is 172.30 pesos or 9 percent larger than under autarky. The difference comes from a combination of a higher level of consumption and a larger utility from completed schooling, as well as from less volatility in household consumption.

5.2 Effects of Conditional Cash Transfer

In the second counterfactual experiment I introduce a conditional cash transfer (CCT) program, which is similar to the design of PROGRESA to both the baseline risk-sharing model and the autarky model. Actual PROGRESA beneficiaries were selected based on multidimensional household characteristics, such as the main materials of walls and floors of household dwellings and ownership of durable goods and assets. Because my model does not have these variables I cannot apply the same criteria. Instead, I choose the low-income type households in the model as beneficiaries. To make the exercise compatible with the actual policy, I take the actual PROGRESA educational grant schedule and convert it into adult-equivalent terms and adjust for the economies of scale as was done for

[23]Under these assumptions, and with the discount factor of 0.95 and risk aversion coefficient of 0.63, the lifetime utility of a new household is given by:

$$\text{(lifetime utility)} = \sum_{t=1}^{120} \delta^{t-1}\left(\frac{C^{1-\eta}-1}{1-\eta}\right) = \frac{1-\delta^{120}}{1-\delta}\left(\frac{C^{1-\eta}-1}{1-\eta}\right) = 19.08\left(\frac{C^{0.37}-1}{0.37}\right),$$

and the consumption-equivalent term, C, of the given lifetime utility is calculated by the following formula:

$$C = \left[\frac{\text{(lifetime utility)}0.37}{19.08}+1\right]^{\frac{1}{0.37}}.$$

Table 18: Educational grant schedule used in counterfactual experiment, 12≤age≤18

School level	Grade	Monthly grant in pesos
Primary	3	145.83
	4	168.48
	5	211.22
	6	269.15
Secondary	1	410.69
	2	403.86
	3	457.14

Table 19: Educational grant schedule used in counterfactual experiment, 6≤age≤11

Age	Monthly grant in pesos
6	0
7	0
8	145.83
9	145.81
10	170.14
11	206.50

household income and consumption.[24] Because households vary in their size, the adjusted amount varies across households. I take the median value of the adjusted amount of program transfers in each grade level and introduce the amount to the model. Because I do not track the school grade of children before age 12 in the model, for that age group it is not feasible to implement the subsidy schedule based on their completed school grade. Instead, I construct the subsidy schedule based on their age instead of school grade. The constructed semi-annual educational grant schedule is provided in Tables 18 and 19. Table 20-Table 23 provide the results of the CCT experiments.

The effect of the CCT on schooling outcomes under risk sharing is shown in Table

[24] To do this, I first assign the program transfers as shown in Table 1 to the sample households. The amount of the program transfers assigned to each household corresponds to the completed school grade of the oldest child of that household. Then I rescale the amount of program transfers for each household by an adult-equivalent household size and adjust for the economies of scale.

20. The fraction of children attending school between ages 12 and 18 among beneficiary households increases by 0.24, from 0.51 to 0.75, and their completed grade at age 19 increases by 1.17 years, from 7.11 to 8.28. There is little change for non-beneficiary households. There is no change in the outside option value for the non-beneficiary households and their demand for risk sharing is not affected by the introduction of the CCT. Also, the current model assumes that a village is populated by a continuum of households. Because idiosyncratic risks can be pooled among each type and any difference in permanent or deterministic components in income are not shared, the high-income type households are not much affected by changes among the low-income type households. Note that under this experiment households fully anticipate the presence of subsidy from the beginning of their lifecycle. Although there is an equilibrium effect, the effect turns out to be very small on the high-income type (non-beneficiary) households.[25]

Table 20: The effect of CCT on schooling outcome under risk sharing

	No CCT		CCT	
	All	Beneficiary	All	Beneficiary
Fraction attending school	0.55	0.51	0.73	0.75
Completed grade at age 19	7.53	7.11	8.37	8.28

Table 21: The effect of CCT on income, consumption, and welfare outcome under risk sharing

	No CCT	CCT
Household consumption	2,634.46	2,661.46
Beneficiary household consumption	1,667.29	1,839.44
Correlation of income and consumption	0.10	0.12
Welfare in consumption equivalence	1,926.66	1,952.93
CCT expenditure	-	28.43

Notes: The unit of the reported numbers is Mexican peso.

[25] Among high-income type households, the fraction attending school slightly increases from 0.6685 to 0.6689 and the completed grade at age 19 increases from 8.5825 to 8.5836. The changes are negligible and are thus not reported in the tables.

Table 22: The effect of CCT on schooling outcome under autarky

	No CCT		CCT	
	All	Beneficiary	All	Beneficiary
Fraction attending school	0.44	0.39	0.69	0.74
Completed grade at age 19	7.10	6.68	8.22	8.26

Table 21 shows the effect of the CCT on household income, consumption, and welfare. The reported outcomes are at the village-aggregate level which includes both beneficiary and non-beneficiary households. The household consumption level with the CCT is higher by 27 pesos, and the welfare in consumption equivalence increases by 26.27 pesos. The correlation between household income and consumption increases by 20 percent, from 0.10 to 0.12. This is because CCT increases the outside option value of beneficiary households which makes (VP) of these households bind more. The reported CCT expenditure is total CCT amount claimed by beneficiary households in all villages divided by the total measure of households. The CCT expenditure computed in this way is 28.43 pesos, which means that the CCT budget was large enough to give 28.43 pesos to every household in the village in each period. The return from the CCT is calculated as the change in the outcome of interests divided by the CCT expenditure. The computed return is 0.95 for household consumption and 0.92 for welfare.

The effect of the CCT on schooling outcomes under autarky is shown in Table 22. Because there is no interaction among households, non-beneficiary households are not affected by the introduction of the CCT. The fraction of children attending school among beneficiary households increases by 0.35, from 0.39 to 0.74. Their completed grade at age 19 increases by 1.58 years, from 6.68 to 8.26. Household consumption and welfare in consumption equivalence increases by 37.10 and 34.36 pesos respectively. The CCT expenditure is 28.05 pesos. The return from the CCT is computed as in under risk sharing, and it is 1.32 and 1.22 respectively for household consumption and welfare.

Table 23: The effect of CCT on income, consumption, and welfare outcome under autarky

	No CCT	CCT
Household consumption	2,629.79	2,666.89
Beneficiary household consumption	1,802.63	2,025.85
Welfare in consumption equivalence	1,770.81	1,805.17
CCT expenditure	-	28.05

Notes: The unit of the reported numbers is Mexican peso.

In comparison to the outcomes under autarky, the effects of the CCT under risk sharing are smaller. Under risk sharing, without the CCT, the fraction of children attending school among beneficiary households is 12 percentage points higher under risk sharing than under autarky; however, with the CCT, the difference is only one percentage point. The returns from the CCT in terms of household consumption and welfare are much larger under autarky. There are two possible explanations. First, the CCT also serves as insurance under autarky. Without the CCT, there are more borrowing-constrained households under autarky than under risk sharing. The CCT not only increases the return from schooling but also relaxes the borrowing constraints of households which are hit by negative parental income shocks. However, under risk sharing, the latter effect is negligible because households already had informal insurance that was close to complete risk sharing. Second, this may be due to the design of the program transfer schedule which gives cash transfer only up to the 9th grade. Because there are more children who are closer to or above the 9th grade under risk sharing than under autarky without CCT, there are more households that would respond to CCT under autarky (“ceiling effect”).

Table 24: (a) Fraction of children attending school and (b) mean household consumption under autarky by cross-sectional parental-income states

	(a) Fraction attending school		
Deviation from the mean log income	No CCT	CCT	Diff
- two s.d.	.36	.68	+.32
- one s.d.	.41	.69	+.28
zero	.44	.70	+.26
+ one s.d.	.47	.68	+.21
+ two s.d.	.49	.64	+.15

	(b) Mean consumption		
Deviation from the mean log income	No CCT	CCT	Diff
- two s.d.	623.03	787.74	+165.02
- one s.d.	1002.22	1167.12	+165.39
zero	1827.51	1988.95	+162.32
+ one s.d.	3741.84	3893.40	+152.61
+ two s.d.	7699.53	7833.98	+136.44

Notes: The unit of the reported numbers is Mexican peso.

Table 25: Effect of CCT with completed grade at age 12 fixed at 2nd grade, beneficiary households only

	Risk sharing			Autarky		
	No CCT	CCT	Diff	No CCT	CCT	Diff
Fraction attending school	0.69	0.94	0.25	0.65	0.93	0.28
Completed grade at age 19	5.90	7.11	1.21	5.75	7.09	1.34

I provide two pieces of evidence which support that the first effect is present. The first piece of evidence is provided in Table 24. The increase in enrollment rate is much larger among the households that were hit by negative income shocks. This suggests that households with greater borrowing constraints were further away from an efficient allocation and responded more strongly to the CCT. The second piece of evidence is obtained by conducting another counterfactual experiment which introduces the same CCT, but under an environment where the cap on the 9th grade does not bind. In this experiment, I force the completed grade at age 12 to be the 2nd grade for every child. As children continue onward from the 2nd grade, the highest grade level they can complete before turning 19 is 9th grade. The schooling outcome under this experiment is provided in Table 25. Again, the differences between the outcomes with and without the CCT in both the fraction attending school and the completed grade level are smaller under risk sharing than under autarky. The results of the original CCT experiment hold even when there is no ceiling effect. However, the difference in the effect of CCT under risk sharing and autarky is smaller in this experiment than in the original CCT experiment, suggesting that ceiling effect is also present.

Another interesting observation from the two counterfactual experiments is that having risk sharing is much more effective in improving the welfare of households than having CCT, although the effectiveness of CCT is greater when it comes to improving schooling outcomes alone. The difference in household welfare under risk sharing and autarky is 1,156.85 pesos (1,770.81 vs. 2,926.66 pesos). When CCT is introduced to autarky economy the welfare increases only by 34.36 pesos from 1,770.81 to 1,805.17 pesos.

6 Conclusion

This paper is the first to incorporate school attendance choices into an inter-household risk-sharing model with limited-commitment constraints. This paper develops and structurally estimates a dynamic risk-sharing model with limited-commitment constraints and school attendance choices in order to study the effectiveness of informal risk sharing in smoothing consumption and schooling, and to evaluate CCT programs when inter-household transfers are allowed. The model considers a village economy comprised of overlapping generations of heterogeneous households. The model parameters are estimated by the simulated method of moments matching the observed income, consumption, and child activity choices predicted by the model to those in data from PROGRESA program in Mexico.

Based on the estimated model, I find that the PROGRESA villages are able to smooth consumption and schooling against idiosyncratic income shocks effectively through inter-household transfers. The amount of consumption and children's school attendance or labor choices are not significantly affected by idiosyncratic income shocks. In contrast to this, a counterfactual simulation of an economy under autarky shows that without inter-household transfers, children's school attendance and labor choices as well as the amount of consumption are substantially affected by income shocks. The estimated utility cost of attending school for children whose schooling was interrupted in the past is substantial. Therefore, the effect of transitory negative income shocks on children's school attendance under autarky accumulates over time, leading to lower schooling outcomes compared to the outcomes under risk sharing. Under autarky, the fraction attending school is 11 percentage points lower than that with risk sharing and the completed school grade by age 19 is higher by 0.47 years.

I also conduct counterfactual policy simulations by introducing a CCT program similar in design to PROGRESA to the model, both with and without risk sharing. Based on the counterfactual simulations I find that the effect of the CCT program on schooling outcomes and the welfare of households is larger under autarky than under risk sharing. The CCT not only increases net returns from schooling by reducing the opportunity cost of attending school, but also mitigates the effect of negative income shocks on household consumption and school attendance. The former effect of the CCT is common under both risk sharing and autarky. The latter effect of the CCT, however, is negligible under risk sharing because the allocations under risk sharing were close to efficient even without the CCT. Moreover, the simulation results show that, under risk sharing, consumption volatility increases by 20 percents after the introduction of the CCT because beneficiary households can rely on the CCT benefit if inter-household transfer becomes unavailable, and this weakens their voluntary participation incentive to remain in the risk-sharing arrangement. This "crowding-out effect" offsets some of the welfare gains from the CCT under risk sharing. Overall, the simulation outcomes suggest that the benefits of CCT may be inflated if the role of inter-household transfers are not taken into account in an evaluation of CCT programs.

Although consumption risk-sharing models considered in previous studies also capture the crowding-out effect of public transfers, the model developed in this paper is the first to study risk sharing and schooling outcomes jointly under CCT programs which are designed to improve schooling outcomes in particular. This cannot be done without a model that explicitly models school attendance choices. Moreover, introducing school attendance choices to a limited-commitment risk-sharing model adds substantial computational challenges, which I overcome by adopting novel theoretical results and computational algorithms (Clausen and Strub, 2013; Fella, 2013). This is an equilibrium model, and the model is simplified in several dimensions (e.g., with a single-child assumption) to keep computational time manageable. Once computation of the model becomes more tractable, developing a richer model that allows for more realistic evaluation of CCT programs in the presence of inter-household transfers would be possible.

There are a number of avenues for future research. First, this paper studies the effect of an existing program design. Ex-ante evaluation of the alternative CCT program designs in the presence of inter-household transfers will also be of interest to both researchers and policymakers. Second, the model considered here assumes a continuum of households, and thus the analysis was confined to large villages. I plan to conducting similar analysis for small villages with a model that assumes a finite number of households as was considered in Ligon, Thomas, and Worrall (2002), Laczó (2008), and Morten (2013).

A Autarky Problem

I set up an autarky problem for five distinctive lifecycle stages: before a birth occurs (Stage 1), after birth and before a child is aged 12 (Stage 2), while the child is aged between 12 and 18 (Stage 3), after the child finishes schooling and before a migration occurs (Stage 4), and finally, after a migration occurs (Stage 5). In autarky, households maximize their present discounted value of lifetime utility given their period budget constraint,

$$c_{i,t} = y^h_{i,t}$$

There is no saving or borrowing. Autarky problem is presented in functional equations, and autarky value functions will be denoted by $U^{Aut}_{L,S}$ where $L = 1, ..., 5$ is a subscript for a lifecycle stage and $S \in fall, spring$ is a subscript for a semester. Households in Stage 3 make a child activity choice. Households in other stages simply consume $y^h_{i,t}$. Parental type is invariant throughout the lifecycle of a household and is omitted from the equations.

- In Stage 5 households' lifetime utility under autarky is given by,

$$U^{Aut}_{5,fall}(s_{i,t}) = u(y^h_{i,t}) + \delta E_{s_{t+1}} \left[U^{Aut}_{5,spring}(s_{i,t+1}) \right]$$

$$U^{Aut}_{5,spring}(s_{i,t}) = u(y^h_{i,t}) + \delta(1-\pi_d) E_{s_{t+1}} \left[U^{Aut}_{5,fall}(s_{i,t+1}) \right]$$

- In Stage 4,

$$U^{Aut}_{4,fall}(s_{i,t}, X^c_i, ctype) = u(y^h_{i,t}) + \delta E_{s_{t+1}} \left[U^{Aut}_{4,spring}(s_{i,t+1}, X^c_i, ctype) \right]$$

$$U^{Aut}_{4,spring}(s_{i,t},X^c_i,\mu^c_i) = u(y^h_{i,t}) + \delta\left\{(1-\pi_m)E_{s_{t+1}|migrate=0}\left[U^{Aut}_{4,fall}(s_{i,t+1},X^c_i,\mu^c_i)\right] + \right.$$

$$\left. \pi_m E_{s_{t+1}|migrate=1}\left[U^{Aut}_{5,fall}(s_{i,t+1})\right]\right\}$$

where $ctype$ is a child type.

- In Stage 3, when the age of a child is a,

$$U^{Aut}_{3,a,fall}(s_{i,t},d^{school}_{i,t-1},X^c_{i,t}) = \max_{d_{i,t}}\left\{u(y^h_{i,t}) + \delta E_{s_{t+1}}\left[U^{Aut}_{3,a,spring}(s_{i,t+1},d^{school}_{i,t},X^c_{i,t})\right]\right\}$$

and in the spring, if $a < 18$,

$$U^{Aut}_{3,a,spring}(s_{i,t},d^{school}_{i,t-1},X^c_{i,t}) =$$
$$\max_{d_{i,t}}\left\{u(y^h_{i,t}) + \delta E_{s_{t+1}}\left[U^{Aut}_{3,a+1,fall}(s_{i,t+1},d^{school}_{i,t},X^c_{i,t+1})\right]\right\}$$

and if $a = 18$,

$$U^{Aut}_{3,a=18,spring}(s_{i,t},d^{school}_{i,t-1},X^c_{i,t}) =$$
$$\max_{d_{i,t}}\left\{u(y^h_{i,t}) + \delta E_{s_{t+1}}\left[U^{Aut}_{4,fall}(s_{i,t+1},X^c_{i,t+1},ctype)\right]\right\}$$

- In Stage 2,

$$U^{Aut}_{2,a,fall}(s_{i,t}) = u(y^h_{i,t}) + \delta E_{s_{t+1}}\left[U^{Aut}_{2,a,spring}(s_{i,t+1})\right]$$

and in the spring, if $a < 11$,

$$U^{Aut}_{2,a,spring}(s_{i,t}) = u(y^h_{i,t}) + \delta E_{s_{t+1}}\left[U^{Aut}_{2,a,fall}(s_{i,t+1})\right]$$

and if $a = 11$,

$$U^{Aut}_{2,a,spring}(s_{i,t}) = u(y^h_{i,t}) + \delta E_{s_{t+1}}\left[U^{Aut}_{3,a+1,fall}(s_{i,t+1}, d^{school}_{i,t} = 1, X^c_{i,t+1})\right]$$

- In Stage 1,

$$U^{Aut}_{1,fall}(s_{i,t}) = u(y^h_{i,t}) + \delta E_{s_{t+1}}\left[U^{Aut}_{1,spring}(s_{i,t+1})\right]$$

$$U^{Aut}_{1,spring}(s_{i,t}) = u(y^h_{i,t}) + \delta\left\{(1-\pi_b)E_{s_{t+1}|birth=0}\left[U^{Aut}_{1,fall}(s_{i,t+1})\right] + \pi_b E_{s_{t+1}|birth=1}\left[U^{Aut}_{0,2,fall}(s_{i,t+1})\right]\right\}$$

B Solution Algorithm for the Agent's Problem

In this subsection, I describe the numerical solution algorithm for the agent's problem. For the computation, the supports of ε^p, ε^h, ε^w, and ξ are discretized. The elements in the discretized supports will be denoted by $\bar{\varepsilon}^p$,$\bar{\varepsilon}^h$, $\bar{\varepsilon}^w$, and $\bar{\xi}$ in the remainder of this section. The minimum and the maximum value of the discretized support of each shock variable are given by -3σ and $+3\sigma$ respectively where σ is the standard deviation of a given variable. To obtain the right-hand side of (VP) constraints, the autarky problems are solved first. Given the autarky value functions, the agent's problem is solved by the following procedure.

1. Discretize the support of promised utility ω for each state, Ω:

   (a) The lower bound of ω grid is given by the corresponding autarky value

   (b) The upper bound is given by the discounted lifetime utility associated with household consumption equivalent to the largest possible y^h in every period and the largest possible $\bar{\xi}$ and X^c values

(c) A uniform grid of $\log\omega$ with the given lower and upper bounds is constructed.

2. Choose an initial guess of R:[26]

   (a) The lower bound of R is given by $(1-\pi_m)^{\frac{1}{2}}$

   (b) The upper bound of R is given by $\frac{1}{\delta(1-\pi_d)}$.

3. Given R, solve the agent's problem for Stage 4 and 5. Iterate value functions until they converge. This is a direct application of Krueger and Perri (2011).

4. Given $V_{4,fall}(\cdot)$, solve Stage 3 problems by using backward induction starting from the spring of age 18.[27]

   (a) Given each choice of d_t, solve the following problem:

$$
\begin{aligned}
& V_{3,a=18,spring}(\omega_t,\Omega_t|d_t) = \\
& \min_{h_t,\{\omega_{t+1}(\Omega_{t+1})\}_{\forall\Omega_{t+1}}} \Big\{ C(h_t,d_t) - w_t d_t^{work} + \\
& \frac{1}{R}\sum_{\Omega_{t+1}} \Pr(\Omega_{t+1}|\Omega_t,d_t) V_{4,fall}(\omega_{t+1}(\Omega_{t+1}),\Omega_{t+1}) \Big\}
\end{aligned}
\tag{6}
$$

   subject to

$$\omega_t = h_t + \delta E_{\Omega_{t+1}|\Omega_t,d_t}\left[\omega_{t+1}(\Omega_{t+1})\right] \ \forall\Omega_{t+1}$$

$$\omega_{t+1}(\Omega_{t+1}) \geq E_{\Omega_{t+1}|\Omega_t,d_t}\left[U_{t+1}^{Aut}(\Omega_{t+1})\right] \ \forall\Omega_{t+1}.$$

[26] The lower and upper bounds of R are the smallest and the largest value with which the discounted sum of the agent's cost is bounded and the convergence of the value functions in Stage 1, 4, and 5 are ensured

[27] I ignore ε^w and ξ throughout this section for the simplicity of the presentation of the algorithm. In practice, these shocks have to be included as state variables whenever needed.

Here,

$$\{\omega_{t+1}(\Omega_{t+1})\}_{\forall\Omega_{t+1}} = \{\omega_{t+1}(\bar{\varepsilon}^h, X, ctype = low), \omega_{t+1}(\bar{\varepsilon}^h, X, ctype = high),$$
$$\omega_{t+1}(\bar{\varepsilon}^h, X+1, ctype = low), \omega_{t+1}(\bar{\varepsilon}^h, X+1, ctype = high)\}_{\forall\bar{\varepsilon}^h}$$

Optimal h_t and $\{\omega_{t+1}(\Omega_{t+1})\}_{\forall\Omega_{t+1}}$ can be found by the first order condition:[28]

$$\frac{\partial V_{4,fall}(\omega^*_{t+1}, \Omega_{t+1})}{\partial \omega_{t+1}} = R\delta \frac{\partial C(h^*_t, d_t)}{\partial h_t}$$

$\omega_{t+1}(\Omega_{t+1}) = \omega^*_{t+1}$ for all Ω_{t+1} that (VP) does not bind, and $\omega_{t+1}(\Omega_{t+1}) = U^{Aut}_{4,fall}(\Omega_{t+1})$ otherwise.[29]

(b) Then, choose d_t that minimizes $V_{3,a=18,spring}(\omega_t, \Omega_t | d_t)$.

5. Solve the agent's problem in the fall of age 18. Given each choice of d_t, the structure of the problem is the same as in (6) but with $V_{3,a=18,spring}(\omega_{t+1}(\Omega_{t+1}), \Omega_{t+1})$ as the continuation value function. Here,

$$\{\omega_{t+1}(\Omega_{t+1})\}_{\forall\Omega_{t+1}} \quad = \quad \{\omega_{t+1}(\bar{\varepsilon}^p)\}_{\forall\bar{\varepsilon}^p}$$

Note that the continuation value function, $V_{3,a=18,spring}(\omega_{t+1}(\Omega_{t+1}), \Omega_{t+1})$, contains a non-differentiable point along the dimension of ω_{t+1} due to a discrete school

[28]In the spring of age 18, the continuation value function, $V_{4,fall}(\cdot)$, is convex and differentiable because the Stage 4 problem does not involve any discrete choice. Thus, the solutions can be found from the first order condition given a choice of d_t. Multiple events occur between the end of age 18 and the beginning of Stage 4: (1) the grade failure or completion is determined; (2) given the final school grade, the child type is drawn; (3) $\bar{\varepsilon}^h$ is drawn. Thus, for each given $\bar{\varepsilon}^h$, there are four different states that the promised utility needs to be assigned.

[29]Because $\bar{\varepsilon}^h$ is i.i.d., as long as (VP) does not bind, the same promised utility value is assigned for all $\bar{\varepsilon}^h$ states. For $\frac{\partial V_{4,fall}(\omega^*_{t+1}, \Omega_{t+1})}{\partial \omega_{t+1}}$ I use the derivative of $C(\cdot)$ at the optimal level of h_{t+1} which is obtained in the fall of Stage 4 (envelope condition). Because ω^*_{t+1} is usually not on the ω_{t+1} grid, I interpolate for an optimal h_{t+1} value which is associated with $(\omega^*_{t+1}, \Omega_{t+1})$.

attendance choice which will be made in the spring of age 18.[30]

(a) Given a choice of d_t, identify a non-convex interval in $V_{3,a=18,Spring}(\omega_{t+1},\cdot)$. Specifically, find the points on the ω_{t+1} grid where the non-convex interval begins and ends.

(b) (Endogenous grid method) For some ω^c_{t+1} which is on the ω_{t+1} grid and is contained in the convex interval, find h^c_t that satisfies the first order condition[31]:

$$\frac{\partial V_{3,a=18,spring}(\omega^c_{t+1},\Omega_{t+1})}{\partial \omega_{t+1}} = R\delta \frac{\partial C(h^c_t,d_t)}{\partial h_t}$$

For Ω_{t+1} in which (VP) does not bind $\omega_{t+1}(\Omega_{t+1}) = \omega^c_{t+1}$. Otherwise,

$$\omega_{t+1}(\Omega_{t+1}) = U^{Aut}_{3,age=18,Spring}(\Omega_{t+1}).$$

Then, back out ω^c_t using (PK),

$$\omega^c_t = h^c_t + \delta \sum_{\Omega_{t+1}} \Pr(\Omega_{t+1}|\Omega_t,d_t)\omega_{t+1}(\Omega_{t+1})$$

Note that ω^c_t may not be a point on the ω_t grid.[32] Compute the cost of delivering ω^c_t, $V_{3,a=18,fall}(\omega^c_t,\Omega_t|d_t)$.

(c) For some point ω^{nc}_{t+1} which is on the ω_{t+1} grid and is contained in the non-convex interval, do the same as above.[33] Back out ω^{nc}_twhich is associated

[30]The continuation value is $E_{\Omega_{t+1|\varepsilon^w,\xi}}[V_{3,a=18,spring}(\omega_{t+1}(\Omega_{t+1}),\Omega_{t+1},\cdot)]$, but I will ignore $\{\varepsilon^w,\xi\}$ as I mentioned in the earlier footnote.

[31]The first order condition given here is well defined because $V_{3,age=18,Spring}$ is differentiable at the solution to this first order condition, ω^*_{t+1}, given a point ω^*_t(Clausen & Strub, 2013).

[32]The reason I use the endogenous grid method instead of solving for a set of optimal $\{\omega_{t+1}\}$ given ω_t is because when numerically solving for $\{\omega_{t+1}\}$, the solver may step into the non-convex interval, in which case the numerical solver does not function well.

[33]The derivative $\frac{\partial V_{3,a=18,spring}(\omega^{nc}_{t+1},\Omega_{t+1})}{\partial \omega_{t+1}}$ is well-defined and the envelope condition also holds because the optimal solution is always found on the smooth part of the objective function (Clausen and Strub, 2013).

with ω_{t+1}^{nc} and compute $V_{3,a=18,fall}(\omega_t^{nc},\Omega_t|d_t)$. However, since ω_{t+1}^{nc} is contained in a non-convex interval, the first order condition is necessary but not sufficient. Thus, given ω_t^{nc}, we need to check whether $\{h_t^{nc},\omega_{t+1}^{nc}\}$ minimizes the cost of delivering ω_t^{nc}.

(d) Given ω_t^{nc} solve the agent's problem using a grid-search method searching over the ω_{t+1} grid. If $\{h_t^{nc},\omega_{t+1}^{nc}\}$ indeed minimizes the agent's cost given ω_t^{nc}, keep the pair $(\omega_t^{nc},V_{3,a=18,fall}(\omega_t^{nc},\Omega_t|d_t))$. Otherwise, discard the pair.

(e) Do this for all the points on the ω_{t+1} grid.

(f) Then, using the pairs, $(\omega_t,V_{3,a=18,fall}(\omega_t,\cdot))$ obtained in the above steps, interpolate $V_{3,a=18,fall}(\omega_t,\Omega_t|d_t)$ over ω_t grid. See Fella (2013) for more details about and applications of this algorithm.

6. Solve the agent's problem in the spring of age 17. Given each choice of d_t, the structure of the problem is the same as in (6) but with $V_{3,a=18,fall}(\omega_{t+1}(\Omega_{t+1}),\Omega_{t+1})$ as the continuation value function. Here,

$$\{\omega_{t+1}(\Omega_{t+1})\}_{\forall\Omega_{t+1}}=\{\omega_{t+1}(\bar{\varepsilon}^p,X_{t+1}=X_t),\omega_{t+1}(\bar{\varepsilon}^p,X_{t+1}=X_t+1)\}_{\forall\bar{\varepsilon}^p}$$

Unfortunately, the endogenous grid method which was used to solve the agent's problem in the fall of age 18 is not adequate here. This is because the method cannot be applied when there are more than one continuous choice variables. To circumvent this issue, I break the problem into three stages, each of which involve only one continuous choice variable.

(a) In the first stage, the agent only decides how much of ω_t should be delivered as the current period utility, h_t, and how much should be postponed to the next period as a form of promised utility, ω_{t+1}:

$$V_{3,age=17,spring}(\omega_t, \Omega_t|d_t) = \min_{h_t, \omega_{t+1}^{(1)}} C(h_t, d_t) + \frac{1}{R} V_{3,age=18,fall}^{(1)}(\omega_{t+1}^{(1)}, \Omega_t|d_t)$$

subject to

$$\omega_t = h_t + \delta \omega_{t+1}^{(1)}$$

$$\omega_{t+1}^{(1)} \geq \sum_{\Omega_{t+1}} \Pr(\Omega_{t+1}|\Omega_t, d_t) U_{3,age=18,Spring}^{Aut}(\Omega_{t+1})$$

where $V_{3,age=18,fall}^{(1)}(\omega_{t+1}^{(1)})$ comes from the second-stage problem which is described below in 6-(b).

(b) In the second-stage problem, the agent allocates $\omega_{t+1}^{(1)}$ across the two states, grade failure ($X_{t+1} = X_t$) or completion ($X_{t+1} = X_t + 1$):

$$V_{3,age=18,fall}^{(1)}(\omega_{t+1}^{(1)}, \Omega_t|d_t) =$$

$$\min_{\omega_{t+1}^{(2)}(X_t), \omega_{t+1}^{(2)}(X_t+1)} \Big\{ \pi_{gr} V_{3,age=18,fall}^{(2)}(\omega_{t+1}^{(2)}(X_t+1), X_{t+1} = X_t + 1, \Omega_t|d_t)$$
$$+ (1-\pi_{gr}) V_{3,age=18,fall}^{(2)}(\omega_{t+1}^{(2)}(X_t), X_{t+1} = X_t, \Omega_t|d_t) \Big\}$$

subject to,

$$\omega_{t+1}^{(1)} = \pi_{gr} \omega_{t+1}^{(2)}(X_t + 1) + (1 - \pi_{gr}) \omega_{t+1}^{(2)}(X_t)$$

$$\omega_{t+1}^{(2)}(X_{t+1}) \geq \sum_{\Omega_{t+1}} \Pr(\Omega_{t+1}|\Omega_t, d_t) U_{3,age=18,spring}^{Aut}(\Omega_{t+1}) \text{ for } X_{t+1} = X_t, X_t + 1$$

where π_{gr} is the probability that the child completes a school grade and $V_{3,age=18,fall}^{(2)}$ comes from the third-stage problem which is described below.

(c) In the third-stage problem, the agent determines the allocation of $\{\omega_{t+1}^{(3)}(\bar{\varepsilon}^p)\}_{\bar{\varepsilon}^p}$

given X_{t+1} and $\omega^{(2)}_{t+1}(X_{t+1})$:

$$V^{(2)}_{3,age=18,fall}(\omega^{(2)}_{t+1}(X_{t+1}),X_{t+1},\Omega_t|d_t) =$$
$$\min_{\{\omega^{(3)}_{t+1}(\bar{\varepsilon}^p)\}_{\bar{\varepsilon}^p}} \sum_{\bar{\varepsilon}^p} \Pr(\Omega_{t+1}|\bar{\varepsilon}^p, X_{t+1}, d_t) V_{3,age=18,fall}(\omega^{(3)}_{t+1}(\bar{\varepsilon}^p), \Omega_{t+1})$$

subject to,

$$\omega^{(2)}_{t+1}(X_{t+1}) = \sum_{\bar{\varepsilon}^p} \Pr(\bar{\varepsilon}^p)\omega^{(3)}_{t+1}(\bar{\varepsilon}^p)$$

$$\omega^{(3)}_{t+1}(\bar{\varepsilon}^p) \geq U^{Aut}_{3,age=18,fall}(\Omega_{t+1}|\bar{\varepsilon}^p, X_{t+1}, d_t) \text{ for all } \bar{\varepsilon}^p$$

(d) I use the endogenous grid method for the first-stage problem and the grid search method for the second- and the third-stage problems. It turns out that a coarse grid in the second- and the third-stage problems leads to failure in finding R that satisfies (RC). Because the problem is solved by backward induction, the third-stage problem is solved first and the first stage the last.

7. Continue to solve the agent's problem to age 0 by backward induction. Then, given the value function at age 0, $V_{2,age=0,fall}(\omega_t)$, solve the Stage 1 problems using the value function iteration.

8. Compute the stationary distribution $\Phi(\Omega)$[34].

   (a) Denote $\Phi^{(0)}_{1,fall}$ to be the initial distribution over the set of $\{\omega_t, \bar{\varepsilon}_t{}^p\}$ in the fall of Stage 1. Obtain $\Phi^{(0)}_{1,spring}$ over $\{\omega_{t+1}, \bar{\varepsilon}^p_{t+1}\}$ by updating ω_{t+1} using the optimal solution found by solving the agent's problem.

   (b) At the end of the spring, households of measure π_b experience a birth and exit to Stage 2. Also, the new households of measure π_b enter the village. Update

[34] $\Phi(\Omega)$ is village specific. The solution of the agent's problem also differs across different villages

$\Phi^{(1)}_{1,fall}$ using the optimal solution. Assign the same $\{\omega_{t+1}, \bar{\varepsilon}^p_{t+1}\}$ distribution to the new households.

(c) Repeat until $\Phi^{(n)}_{1,fall} \approx \Phi^{(n+1)}_{1,fall}$. For $(1-\pi_m)^{\frac{1}{2}} < R < \frac{1}{\delta(1-\pi_d)}$, there exists a unique stationary distribution $\Phi^*_{1,fall}$ (Krueger and Perri, 2011).

(d) Given $\Phi^*_{1,fall}$ the distribution over the households in different life-cycle stages is computed by updating the distribution using the optimal solution of the agent's problem.

9. Compute the excess demand and update R.

    (a) Simulate household consumption and income and compute the excess demand (village aggregate consumption - village aggregate income including children's wage income) using the stationary distribution.

    (b) If the excess demand is positive, decrease R. If it is negative, increase R. I use a bisection method.

10. Repeat from step 1 to step 9 until the excess demand becomes close to zero.

C Household Consumption and Income Variables

C.1 Consumption

Consumption measures are consistent throughout the three data waves of October 1998, May 1999, and November 1999. The data contains detailed information on food consumption and non-food expenditures by category.

Households report their weekly food consumption during the week prior to the survey on a recall basis. For each item, they report the quantity consumed, quantity purchased, expenditure on the purchase, and the quantity of home-produced goods that were consumed within the household.[35] I obtain the unit price of each item by dividing the expenditure on the purchase by the quantity purchased. Then, the weekly food consumption variable was constructed by multiplying the unit price and the quantity consumed (including the home-produced goods) for each item and then adding up the values for all food items. When the expenditure variable is missing so that the actual unit price paid by a household is not available, I use the village median price if there are more than 20 households through which the unit price could be computed. If there are fewer than 20 households providing available prices, I use the municipal median price; if the number of the available prices in a municipal is less than 20, the province median price is used. Finally, the semiannual household food consumption variable is constructed by multiplying the weekly food consumption variable by 26.

The non-food expenditure category includes transportation in the past week; electricity, fuel, hygienic products, medication, and school fees in the past month; and household utensils, home and personal accessories, clothing, shoes, toys, transportation, school supplies, ceremonies, and expenditures on home improvement in the past six months. I convert the weekly expenditure and monthly expenditures to a semiannual expenditure by multiplying by 26 and 6 respectively. The non-food expenditure variable is constructed as the sum of semiannual expenditures on the all non-food categories.

Household consumption is the sum of the food consumption and non-food expenditure variables. Inflation is adjusted by the monthly CPI. I use the price in October 1998 as a base; May 1999 and November 1999 values were deflated by 1.10998 and 1.1392 respectively.

[35]Items in the food consumption survey are the following. The "fruit and vegetable" category includes tomatoes, onions, potatoes, carrots, leafy vegetables, oranges, bananas, apples, lemons, and pears. The "grain" category includes corn, maize, bread, flour, pasta, rice, cookies, and beans. The "meat and dairy" includes chicken, beef, pork, sheep, goat, fish, seafood, canned fish, eggs, milk, and cheese. The "industry goods" category includes coffee, sugar, alcohol, vegetable oil, and refreshments.

Table 26: Income by source (in 1998 pesos)

Source	October 1998	May 1999	November 1999
Wage	5281.5 (7399.3)	3984.6 (5420.1)	5154.2 (6479.0)
Crop production	1987.0 (8001.8)	1022.4 (5021.7)	-
Livestock	42.8 (750.6)	23.2 (524.2)	-
Sales of service	175.8 (1932.0)	75.5 (744.2)	32.3 (463.8)
Public transfer	354.0 (956.4)	295.6 (799.6)	259.5 (779.1)
Other	90.98 (946.0)	57.9 (657.6)	760.5 (3903.0)
Observations	7,405	5,307	5,661

Source: PROGRESA data from October 1998, May 1999, and November 1999. Control villages data only. Other income includes pensionss, interest payment, rent, and income from community gains. The number of observations indicates the number of households that had non-missing observations for all income categories.

C.2 Income

Household income is constructed by finding the sum of the wage income of household members, the crop production and livestock income, and the income from other sources such as rent, public program transfers, and the sales of the services of household members. Table 26 provides the mean and the standard deviation of income from different sources.

In the wage income category, for every household member above age 6, the number of days worked in the past week of the survey, the number of weeks worked in the past month, and the amount and frequency of the wages received are reported. The wage rate is calculated using the information about the amount of wages received and the frequency of the payment. The number of days worked in the past week and the number of weeks worked in the past month are also available. Using the calculated wage rate and the days and weeks worked, the monthly wage rate is computed. For example, when the weekly wage rate is available for a given individual, and the individual worked for 3 weeks in the past week, then the monthly wage income is the product of the weekly wage rate multiplied by 3 weeks. Monthly wage income is then multiplied by 6 and converted into a semiannual wage income.

The October 1998 and May 1999 waves also contain detailed information about crop and livestock production. For each plot used for crop cultivation, the type of crop, quantity sold, quantity harvested, income from the crop sales, and the expenditure on inputs such as fertilizer, seeds, insecticides, machinery, and labor are reported. Crop production income is computed by subtracting the total expenditure on production inputs from the value of the total quantity harvested. The value of the harvest for each type of crop is calculated by multiplying the unit price by the quantity harvested. I use the producer prices published by the Food and Agriculture Organization of the United Nations for the corresponding years of 1998 and 1999. (I use the value of harvest rather than the sales value because sales do not account for the amount that is unsold and instead left for consumption.) I impute the missing observations for the expenditures on inputs by running a regression of the observed expenditures on the characteristics of the production, such as the type of crop, size of the cultivation area, and municipal dummies. The regression result is reported in Table 27. For those households which did not own any plots or did not use any plots for cultivation, income from crop production is set to zero.

The November 1999 wave did not survey income from crop and livestock production. Instead, the survey included another subcategory for income from the sales of product in other income category. As shown in Table 26, the amount of income in other income category is substantially larger in November 1999 compared to that of October 1998 and May 1999 waves although the amount is smaller than the income from crop and livestock production in the two waves. This may be because the income from the sales of product does not include the value for the produce consumed within the household.

Parental income is the total household income minus the sum of the child wage income.

Table 27: Regression of the log expenditures on crop production inputs

	October 1998		May 1999	
Dependent variable	Coefficient	(s.e.)	Coefficient	s.e.
total cultivation area	0.0116	(0.0124)	0.0516	(0.0421)
irrigated area	0.0231	(0.0657)	-0.0376	(0.0658)
family owned area	0.0367	(0.0028)	0.0727	(0.0045)
rent area	-0.0178	(0.0159)	0.0439	(0.0357)
corn	0.1009	(0.0633)	0.1805	(0.0688)
corn area	-0.0111	(0.0124)	-0.0517	(0.0142)
corn irrigated area	0.0089	(0.0635)	0.0320	(0.0648)
beans	0.2646	(0.0275)	0.2039	(0.0358)
beans area	-0.0004	(0.0005)	-0.0002	(0.0008)
beans irrigated area	0.0188	(0.0233)	0.0575	(0.0320)
tomato	0.6425	(0.1803)	0.3655	(0.3006)
tomato area	-0.0012	(0.0011)	-0.0532	(0.1011)
tomato irrigated area	0.1158	(0.1573)	0.3126	(0.2150)
pumpkin	-0.0440	(0.1407)	-0.0072	(0.1262)
pumpkin area	0.0309	(0.0275)	0.0054	(0.0148)
pumpkin irrigated area	-0.0352	(0.0908)	-0.3145	(0.3359)
chickpea	0.0499	(0.3406)	0.9729	(0.7593)
chickpea area	-0.0295	(0.0613)	-0.2254	(0.2353)
chickpea irrigated area	0.2620	(0.2043)	-0.5883	(0.9868)
wheat	0.0930	(0.3222)	0.3263	(0.2836)
wheat area	0.1417	(0.0877)	-0.0761	(0.0444)
wheat irrigated area	-0.0943	(0.2213)	0.0983	(0.1650)
rice	0.1713	(0.7023)	-0.0623	(2.2196)
rice area	0.0265	(0.1312)	-0.0011	(0.6745)
rice irrigated area	0.2569	(0.2841)	-	-
coffee	0.4191	(0.0745)	0.5073	(0.0754)
coffee area	-0.0150	(0.0130)	0.0239	(0.0160)
coffee irrigated area	0.3143	(0.3383)	-0.0616	(0.0689)

continued

fruit	0.5025	(0.1458)	0.7269	(0.0986)
fruit area	0.0046	(0.0319)	-0.0647	(0.0103)
fruit irrigated area	0.1268	(0.0842)	0.1200	(0.0665)
sugarcane	1.2394	(0.0995)	1.3903	(0.1250)
sugarcane area	0.0303	(0.0131)	-0.0498	(0.0173)
sugarcane irrigated area	0.2052	(0.0973)	0.1683	(0.1233)
chilli	0.3088	(0.2483)	0.5370	(0.2038)
chilli area	-0.0543	(0.0783)	-0.0225	(0.0362)
chilli irrigated area	-0.0670	(0.1651)	0.0917	(0.0756)
potato	0.5793	(0.1997)	0.3508	(0.2434)
potato area	-0.0065	(0.0223)	0.0187	(0.0503)
potato irrigated area	0.6545	(0.2128)	0.0679	(0.1850)
other crops	0.5254	(0.0540)	0.3686	(0.0838)
other crops area	0.0005	(0.0010)	-0.0059	(0.0148)
other crops irrigated area	0.0063	(0.0330)	0.0389	(0.0657)
municipal fixed effect	yes		yes	
R-squared	0.3249		0.3137	
Observations	11,273		7,794	

Source: PROGRESA data from October 1998, May 1999, and November 1999. Both treatment and control villages data are used. Standard errors are in the parenthesis. For example, "Corn" is a 0-or-1 indicator variable that shows whether a household cultivated corn during the period. "Corn area" is the area of the plot that is used for the cultivation of corn. "Corn irrigated area" is the area of the irrigated plot that is used for corn.

References

[1] Albarran, Pedro, and Orazio P. Attanasio. "Limited commitment and crowding out of private transfers: evidence from a randomised experiment." *The Economic Journal* 113.486 (2003): C77-C85.

[2] Atkeson, Andrew, and Robert E. Lucas. "On efficient distribution with private information." *The Review of Economic Studies* 59.3 (1992): 427-453.

[3] Atkeson, Andrew, and Jr, Robert E. Lucas. "Efficiency and equality in a simple model of efficient unemployment insurance." J*ournal of Economic Theory* 66.1 (1995): 64-88.

[4] Angelucci, Manuela, and Giacomo De Giorgi. "Indirect Effects of an Aid Program: How Do Cash Transfers Affect Ineligibles' Consumption?." *The American Economic Review* (2009): 486-508.

[5] Angelucci, Manuela, et al. "Family networks and school enrolment: Evidence from a randomized social experiment." *Journal of Public Economics* 94.3 (2010): 197-221.

[6] Angelucci, Manuela, Giacomo De Giorgi, and Imran Rasul. "Resource Pooling Within Family Networks: Insurance and Investment." Working paper, University of Michigan, 2012.

[7] Attanasio, Orazio P., Costas Meghir, and Ana Santiago. "Education choices in Mexico: using a structural model and a randomized experiment to evaluate Progresa." *The Review of Economic Studies* 79.1 (2012): 37-66.

[8] Attanasio, Orazio, and José-Vıctor Rıos-Rull. "Consumption smoothing in island economies: Can public insurance reduce welfare?." *European Economic Review* 44.7 (2000): 1225-1258.

[9] Behrman, Jere R., Piyali Sengupta, and Petra Todd. "Progressing through PROGRESA: An impact assessment of a school subsidy experiment in rural Mexico." *Economic Development and Cultural Change* 54.1 (2005): 237-275.

[10] Calusen, Andrew, and Strub, Carlo. "A General and Intuitive Envelope Theorem." Working paper, University of Edinburgh, 2013.

[11] Cox, Donald, and Marcel Fafchamps. "Extended family and kinship networks: economic insights and evolutionary directions." *Handbook of Development Economics* 4 (2007): 3711-3784.

[12] Dercon, Stefan, and Pramila Krishnan. "Risk sharing and public transfers." *The Economic Journal* 113.486 (2003): C86-C94.

[13] De Janvry, Alain, et al. "Can conditional cash transfer programs serve as safety nets in keeping children at school and from working when exposed to shocks?." *Journal of Development Economics* 79.2 (2006): 349-373.

[14] Di Maro, "Evaluation of the Impact of PROGRESA on Nutrition: Theory, Econometric Methods and an Approach to Deriving Individual Welfare Findings from Household Data." Mimeo, University College London, 2004.

[15] Fafchamps, Marcel. "Risk sharing between households." *Handbook of Social Economics* (2008): 1-42.

[16] Fafchamps, Marcel, and Susan Lund. "Risk-sharing networks in rural Philippines." *Journal of Development Economics* 71.2 (2003): 261-287.

[17] Fiszbein, Ariel and Norbert Rüdiger Schady. *Conditional cash transfers: reducing present and future poverty*. World Bank Publications, 2009.

[18] Giulio Fella, "A Generalized Endogenous Grid Method for Non-Smooth and Non-Concave Problems", forthcoming *The Review of Economic Dynamics*, 2013.

[19] Jacoby, Hanan G. "Borrowing constraints and progress through school: evidence from Peru." *The Review of Economics and Statistics* (1994): 151-160.

[20] Jacoby, Hanan G., and Emmanuel Skoufias. "Risk, financial markets, and human capital in a developing country." *The Review of Economic Studies* 64.3 (1997): 311-335.

[21] Jensen, Robert T. "Do private transfers 'displace'the benefits of public transfers? Evidence from South Africa." *Journal of Public Economics* 88.1 (2004): 89-112.

[22] Krueger, Dirk, and Fabrizio Perri. "Does income inequality lead to consumption inequality? Evidence and theory." *The Review of Economic Studies* 73.1 (2006): 163-193.

[23] Krueger, Dirk, and Fabrizio Perri. "Public versus private risk sharing." *Journal of Economic Theory* 146.3 (2011): 920-956.

[24] Kinnan, Cynthia. "Distinguishing barriers to insurance in Thai villages." Working paper, Northwestern University, 2011.

[25] Lazcó, Sarolta. "On Risk Sharing in Village Economies: Structural Estimation and Testing." Working paper, 2011.

[26] Lee, Donghoon, and Kenneth I. Wolpin. "Accounting for wage and employment changes in the US from 1968–2000: A dynamic model of labor market equilibrium." *Journal of Econometrics* 156.1 (2010): 68-85.

[27] Ligon, Ethan, Jonathan P. Thomas, and Tim Worrall. "Mutual insurance, individual savings, and limited commitment." *Review of Economic Dynamics* 3.2 (2000): 216-246.

[28] Ligon, Ethan, Jonathan P. Thomas, and Tim Worrall. "Informal insurance arrangements with limited commitment: Theory and evidence from village economies." *The Review of Economic Studies* 69.1 (2002): 209-244.

[29] Lucas, Robert, Nancy Stokey, and Edward Prescott. "Recursive methods in economic dynamics." Cambridge MA (1989).

[30] Mazzocco, Maurizio, and Shiv Saini. "Testing Efficient Risk Sharing with Heterogeneous Risk Preferences." *The American Economic Review* (2012), 428-68.

[31] Morduch, Jonathan. "Income smoothing and consumption smoothing." *The Journal of Economic Perspectives* 9.3 (1995): 103-114.

[32] Morten, Melanie. "Temporary migration and endogenous risk sharing in village india". Working Paper, Stanford University, 2013.

[33] Paul Schultz, T. "School subsidies for the poor: evaluating the Mexican Progresa poverty program." *Journal of Development Economics* 74.1 (2004): 199-250.

[34] Skoufias, Emmanuel. "PROGRESA and its impacts on the welfare of rural households in Mexico." Vol. 139. International Food Policy Research Institute (2005).

[35] Skoufias, Emmanuel, Benjamin Davis, and Jere Behrman. "An evaluation of the selection of beneficiary households in the education, health, and nutrition program (PROGRESA) of Mexico." International Food Policy Research Institute (1999).

[36] Todd, Petra E., and Kenneth I. Wolpin. "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." *The American economic review* (2006): 1384-1417.

[37] Townsend, Robert M. "Risk and insurance in village India." *Econometrica* (1994): 539-591.