Household Responses to Cash Transfers

Denni Tommasi*

ECARES, Université libre de Bruxelles

January 31, 2016

Abstract

This paper estimates a collective model of the household and investigates how parents reach decisions to allocate household resources. Using data from the Mexican PROGRESA, we test the restrictions of collective rationality on a large variety of specifications and show that this modelling approach cannot rationalize the decision process. We provide some evidence that the observed inefficiency is driven by the group receiving the cash transfers. These results are consistent with the idea that a possible (negative) indirect effect of CCT programs may be to enhance disagreements between the spouses which trigger an inefficient allocation of their resources.

JEL Codes: D13, I38, J12, J16, O15

Keywords: collective model, bargaining power, commitment, efficiency, PROGRESA, conditional cash transfers.

*ECARES, Université Libre de Bruxelles. Avenue F. D. Roosevelt 50, CP 114, B-1050 Brussels, Belgium. E-mail: dtommasi@ulb.ac.be. I have benefited from the comments of Arthur Lewbel and Bram De Rock. I thank Gani Aldashev, Samuel Bazzi, Dilip Mookherjee, Tom Potoms, Alexander Wolf, and seminar participants at Boston College, Boston University and ECARES for useful comments and suggestions. I acknowledges financial support from the Fonds National de la Recherche Scientifique (FNRS). All remaining errors are mine.

All models are wrong, but some are useful.
— Box, G. E. P. (1976).

1 Introduction

The success of policies aimed at fighting poverty depends crucially on how parents respond to monetary incentives. If they allocate resources inefficiently, the resulting level of well-being is likely to fall behind the socially efficient optimum. This is undesirable because over the last two decades many policy interventions in developing and developed countries have been channeled through conditional cash transfer (CCT) programs, which have occupied a large percentage of governments' annual anti-poverty budgets. Although there is evidence that they have been beneficial, their effectiveness may still be limited[1].

The collective model of the household, pioneered by Chiappori (1988; 1992) and Apps and Rees (1988) and subsequently elaborated by Browning et al. (1994), Browning and Chiappori (1998), Blundell et al. (2005) and Chiappori and Ekeland (2006), has become the main paradigm through which household allocation decisions are now studied. There are two main reasons for this, which altogether make the framework suitable to study the distributional impacts of public policies. First, the fundamentals of the model, such as individual preferences and the decision process, can be partly identified under reasonable conditions (Chiappori and Ekeland, 2009). Second, the model is based on a small set of assumptions, mainly the Pareto efficiency of the household allocation process, and yet still provides strong testable restrictions. Concerning its main hypothesis, the literature offers various strategies to test collective rationality, mostly based on the use of *distribution factors*, which are variables that affect Pareto weights without changing preferences. The most powerful approach is based on the estimate of a z-conditional demand system, recently introduced by Bourguignon et al. (2009)[2].

Over the last two decades, the collective model has been subject of severe scrutiny, particularly in relation to the efficiency hypothesis. Table 1 in the Appendix summarizes the most important studies testing the efficiency hypothesis within the family. The

[1]This issue is quite relevant if we think, for example, that the main rationale for implementing such programs was the promotion of children's educational and health investments. Inefficient or suboptimal decisions that affect early investments in human capital do not only have an immediate negative impact, but most importantly they are going to have repercussions in later stages of children's lives. For an extensive overview of the causal link between early investment in human capital and long lasting effects, see Cunha et al. (2006).

[2]The other parametric approaches are the proportionality condition test, put forward by Bourguignon et al. (1993) and Browning et al. (1994), which states that if the ratio of of the marginal effect of two distribution factors is equal across demand equations, this is a necessary and sufficient condition for collective rationality. Equivalently, the rank condition test requires that the impact of the distribution factors must be at most of size one (Chiappori and Ekeland, 2006). A completely different, and non-parametric, approach to test the collective model is based on revealed preference theory. In these studies, collective rationality is generally not rejected by the data (Cherchye et al, 2009).

results based on the z-conditional test are potentially the most powerful ones because they boil down to testing single equation exclusion restrictions. However, as it is clear from this table, the collective model is hardly ever rejected, both with z-conditional tests and with other type of tests as well. This under-rejection of the efficiency hypothesis has been recently criticized by Dauphin et al. (2015). In relation to the z-conditional test, they argue that if we apply Bourguignon et al. (2009) strictly, the test requires that at least one distribution factor (locally) affects each demand equation. However, this assumption is hardly ever satisfied empirically, which is problematic because it makes the test procedure relatively weak[3]. Hence, such inconsistency between the theory and the statistical practice, they argue, may be a plausible candidate explaining under-rejection of the collective model[4]. Based on this criticism, they propose a generalization of the z-conditional test that does not require a distribution factor to (locally) affect each demand equation. The basic intuition of this paper is that, even under collective rationality, it is possible that the demand for a subset of goods are not affected by the relative bargaining power of the members, at least locally. In this case, the expenditures on these goods will not depend on distribution factors. Whereas, those goods that are actually affected by the members' bargaining power must depend on all the distribution factors. This "all or nothing" restriction summarizes their necessary condition of efficiency and makes the statistical procedure more easily applicable.

In the present paper we exploit the experimental set-up of a conditional cash transfers (CCT) program in rural Mexico to investigate whether households use their resources efficiently. This dataset has been used recently by some authors for a similar purpose obtaining mixed results (Bobonis, 2009; Attanasio and Lechene, 2014; Angelucci and Garlik, 2015). Our aim is to give a definite answer to this question that is, at the same time, embedded in a general collective framework, theoretically consistent and robust to several specifications. The use of data collected from these experiments is motivated by the fact that randomized programs allow researchers to construct valid distribution factors through which it is possible to test hypothesis about the decision making process of non-unitary households. PROGRESA is a program where the exogenous cash transfers are targeted explicitly to woman, with the objective of deliberately changing the control of resources in the household, increasing the share of total non-labor income controlled by them. What makes PROGRESA unique is the random assignment of the transfers

[3]In practical terms, it means that we end up relaying on the estimate of a parameter that is obtained dividing two numbers that are very small.

[4]There are also other authors who have raised concerns about the fact that the collective model is hardly ever rejected when using tests based on distribution factors. Aguero (2008) points out that the proportionality condition implies a non-linear restriction across equations which is generally tested by a means of a Wald test. However, Wald tests are not invariant to algebraically equivalent non-linear parametrization of the null hypothesis. More recently, Naidoo (2015) question the restrictive nature of the constraints imposed by the efficiency hypothesis as compared with those of a non-cooperative (Nash) approach.

and the richness of the data collected.

To achieve our objective, we first build a dynamic collective model of the household within which we derive the testable implications. Notice that in a dynamic setup we need to distinguish between two concepts of efficiency: ex-ante and ex-post efficiency. "Ex-ante" means that household members commit themselves to future allocations of resources which are Pareto optimal irrespective of the exogenous events that may occur. "Ex-post" means that, after the realization of an event that triggers a renegotiation of the contract, spouses are able to reach a Pareto efficient division of their resources. In a recent compendium of the collective static and dynamic models of the household, Chiappori and Mazzocco (2015) develop a general framework of household behavior and clarify at what stage of the household decision process we can test both concepts of efficiency. The latter is the main reference that we use to derive the empirical tests that are performed in the present paper. For convenience, from now on we refer to the former as *test of commitment* and to the latter as *test of efficiency*. We run both tests after estimating a structural QAIDS model a-la Banks et al. (1997) on household's budget shares of food. Focusing on the budget structure of food is a meaningful exercise as the budget allocated to it accounts for around 80% of expenditure of the households in the sample.

Armed with a theoretically consistent representation of household's preferences for food consumption, we proceed with the analysis as follows. First, we augment the demand system with a distribution factor that was unexpected at the time of marriage formation. This allows us to perform the test of commitment. Controlling for total resources (including those coming from the program), the natural distribution factor to use in order to test commitment is the treatment indicator. Notice that this has been used in the literature and interpreted, in a static setting, as a test of the unitary model (e.g. Attanasio and Lechene, 2002; Rubalcava et al., 2009). Econometrically, there is no difference in results with the present paper because, as it will be clearer in the theoretical section, a rejection of commitment is also a rejection of the unitary model. However, theoretically a test on household behavior using this kind of distribution factor is more appropriate to interpret it as a test of commitment. We perform this test for the full sample and for several sub-samples, with the aim to investigate within-sample and across-time variations. The first set of results indicates that the cash transfers is able to break the commitment between spouses. However, we show that commitment is very strong in the sample of households at our disposal, because it requires a large effect and a high statistical power to be picked up. In light of the theoretical model that guides the empirical analysis, we can interpret these results as the households being inconsistent with respect to the Full-Commitment Intertemporal Collective (FIC) model (Mazzocco, 2007). That is, the exogenous shock that changes the control of resources in the household is able to trigger a renegotiation of the plan of resource allocation.

Second, we proceed our investigation testing the efficiency hypothesis, which represents the core of our analysis, by running both the "all or nothing" restriction test (Dauphin et al., 2015) and the z-conditional demand test (Bourguignon et al., 2009). To make our conclusions more robust, we use three valid distribution factors and focus on all possible pairwise combinations of these variables. The three distribution factors that we use are the treatment indicator, which is exogenous by construction; a variable that measures the relative bargaining strength of the husband and wife within household by using data on the network of relatives present in the village; and a variable of relative education difference between spouses which captures some cultural aspects within the family. The first subset of results is strongly in contrast with the previous literature. On one side, the results of the "all or nothing" restriction indicate that there is no combination of two distribution factors where either both are significant or non significant. On the other side, following the z-conditional test alone lead us instead to a different story. To run this test, we take the demand equations where the distribution factors are most responsive. This is required to increase the power of the test. In this case the results are in favor of collective rationality in 1998, but not in 1999. And we have some ambiguity in the pooled sample. However, in light of the theoretical discussion of Dauphin et al. (2015), we conclude that the collective model is to be rejected in both 1998, 1999 and the pooled sample.

Next, we investigate further whether the group receiving the transfers is driving the observed inefficiency, that is, whether eligibility to PROGRESA may affect the household allocation process. We split the sample between treatment and control groups and run the efficiency test separately for both groups. According to the results of the "all or nothing" restriction for the control group, households behave efficiently in 1998, whereas this hypothesis is harder to defend for the 1999 and the pooled sample. The results of the z-conditional test are instead all in support of the efficiency hypothesis for this group. As for the treatment group, the "all or nothing" restriction rejects collective rationality throughout the period, whereas the z-conditional test alone would not detect inefficiency neither in 1998, 1999 or the pooled sample. All in all, following the Dauphin et al. (2015) sequential approach, there seems to be some indirect evidence, at least for the 1998 data, that participating into the program may drive the overall inefficient allocation of the resources by the household. Unfortunately, our theoretical framework does not allow us to test directly why this might be the case for our sample, whether this is a common behavior when commitment is first broken, or whether it is the particular money of the cash transfers that induces this behavior.

This paper is linked to two streams of literature. The first is the literature investigating whether the ex-ante Pareto efficiency hypothesis holds in households' data. Both Mazzocco (2007), Blau and Goodstein (2014) and Voena (2015) reject this hypothesis. Except Attanasio and Lechene (2002), there is no paper that presents the results of a test on commitment using randomized experiments. With respect to this literature, we show that commitment is difficult to break and requires high statistical power to be picked up. The second is the literature testing the ex-post Pareto efficiency hypothesis. The question of whether households make ex-post efficient decisions has been long studied in the literature. Early papers find efficiency in commodity demand (Bourguignon et al., 1993; Browning et al., 1994; Browning and Chiappori, 1998), labor supply for childless couples (Chiappori et al., 2002; Vermeulen, 2005), demand of children's health (Thomas et al., 2002), time allocation (Bayudan, 2006), female labor supply (Donni, 2007; Donni and Moreau, 2007), demand of child welfare (LaFave and Thomas, 2013) and demand for health insurance (Adamowicz et al., 2013). However, efficiency has been rejected in household agricultural production (Udry, 1996), labor supply for couples with children (Fortin and Lacroix, 1997) and risk sharing activities (Dercon and Krishnan, 2000; Robinson, 2012).

More recently, and using the same dataset used in the present paper, Bobonis (2009) exploits the random assignment of PROGRESA in Mexico, and rainfall shocks during the period of the evaluation surveys, and show that parents make efficient decisions with respect to household demand for general commodities. Attanasio and Lechene (2014) use the same dataset, but combine it with kinship networks within the village, and show efficiency of household demand for food items. Angelucci and Garlik (2015) follow the previous authors and present evidence of within-sample variation in the efficiency of intra-household resource allocation. They observe that the consumption patterns of these rural households are Pareto efficient when the heads are relatively old, but not when they are relatively young. With respect to this literature, particularly the last three papers cited, we obtain results that are in strong contrast, meaning that we reject the efficiency hypothesis, albeit we use the same dataset. The different results are due to two main reasons. First and foremost, we use a more general test that makes the empirical procedure more easily applicable and the results more robust. Second, we use a specification of the demand system that follows more closely the traditional QAIDS a-la Banks et al. (1997). As it will be discussed in Section 5.3, the previous literature testing collective rationality on PROGRESA data relies on the estimates of a linear version of the QAIDS which is likely to be problematic. Estimating a proper QAIDS is feasible to do as we have precise information on quantities and prices for several commodities. Moreover, differently from the previous literature, we provide some evidence that the inefficiency may be driven by the participation into the conditional cash transfers program.

The remainder of this paper is organized as follows. In Section 2 we describe the theoretical framework that motivates the empirical analysis. In Section 3 we discuss the data of the welfare program that we use. In section 4 we discuss the empirical strategy and the methodological issues related to the estimation of a demand system. In section 5, we present the empirical results and the robustness checks. Section 6 concludes.

2 Theoretical framework

In this section, we discuss the theoretical set-up of individuals' interactions within the household. Following Chiappori and Mazzocco (2015), we develop a simple version of the Limited-commitment Intertemporal Collective (LIC) model to motivate our empirical analysis[5]. The main feature of the LIC model is that household decisions are efficient subject to the constraint that in each period and state of nature both spouses can choose to leave the household. This means that both members value the level of welfare provided by staying in the household at the current allocation of resources, vis-a-vis the best outside option, and choose which option to take. In such a context, the definition of the outside option is crucial. Here we mimic the description of the threat point provided by Lundberg and Pollak (1993), which is defined as the individual welfare if the two spouses choose not to cooperate within marriage[6]. The reminder of this section is organized as follows: first, we present the basic household dynamic set-up; second, we re-interpret it as a three-stage formulation following Chiappori and Mazzocco (2015) and we derive the empirical predictions for each stage that will be tested in the empirical section.

2.1 Preferences and decisions

Consider a household comprising two decision makers $i \in \{m, f\}$ and any number of children, where m stands for mother, f for father, and where children are not part of the decision making process and enter as a public good within the household. There is a 2-period horizon. In the first period the household is formed, which implies that the lifetime expected value of the match exceeds the value of the outside option for both spouses. In the second period a new policy is installed which bears two important characteristics: first, a (randomly chosen) fraction of households is eligible to receive a large conditional cash transfers and, second, eligibility affects the outside option of member m *only*. This means that member m is entitled of the cash transfers which will lead to an increase of her outside option.

Household member i in period t cares about her own private consumption $\mathbf{q}_{it}$ and a household public good $\mathbf{Q}_t$. For simplicity we assume that leisure is separable as well as utility over time, there is no discounting and households cannot save from one period to another[7]. Each member's preferences are assumed to be representable by a continuously

[5]For a general version of the model, see Ligon (2002), Aura (2005), Mazzocco (2007), Gemici and Laufer (2012), Mazzocco, Ruiz, and Yamaguchi (2014), Lise and Yamada (2014), Voena (2015), or Bronson (2015).

[6]In the literature it is possible to find also an alternative definition which is compatible with the non-cooperative behavior idea: the outside option is defined as the value of being divorced. Note that the threat to divorce does not necessarily imply that the household members will actually pursue such an option. In many cases, hitting the threat point will trigger a renegotiation that modifies the intra-household allocation plans to make the new plans at least as good as the outside option for each member.

[7]This is not going to be a restrictive assumption in our context as the empirical application is

differentiable and strictly concave utility map $u_i(\mathbf{q}_{it}, \mathbf{Q}_t)$. The extent to which members m and f care about the children is captured by their preferences for the public good. The resources of the family are derived from a household-level endowment A_1, realized at the time of marriage, total household earnings x_t, realized at each point in time, and an endowment entitled to member m which is realized in period 2. The endowment is realized at the beginning of period 2, before period-2 choices are made. This cash transfers is the only source of uncertainty in the model. The state-contingent budget constraint of the family in each period can then be written as follows:

$$\mathbf{p}_t'\mathbf{q}_t + \mathbf{P}_t'\mathbf{Q}_t = A_1 + x_t + \mathbb{1}(t = 2) \cdot z_2^* \tag{1}$$

where $\mathbf{p}_t$ and $\mathbf{P}_t$ are the price vectors of private and public goods respectively, and z_2^* is the realization of the random variable z_2, which is the cash transfers entitled to member m in period 2.

Pareto efficiency implies that households choose how to allocate their income to consumption by maximizing a weighted sum of the spouses' expected value of lifetime utility:

$$\max_{\{\mathbf{q}_{mt}, \mathbf{q}_{ft}, \mathbf{Q}_t\}} \sum_t \sum_i E_0 \big[M_{it} \cdot u_i(\mathbf{q}_{it}, \mathbf{Q}_t) - \lambda_{it}(\mathbf{Z}) \cdot \bar{u}_{it}(\mathbf{Z}) \big] \tag{2}$$

subject to (1), where E_0 is the expectation operator, $M_{it}(\mathbf{Z})$ is the Pareto weight (or bargaining power) of member i at time t, $\lambda_{it}(\mathbf{Z})$ is the Lagrangian multiplier corresponding to the participation constraint of member i at time t, $\bar{u}_{it}(\mathbf{Z})$ denotes the value of the outside option for member i and time t, and $\mathbf{Z}$ is a set of *distribution factors*. Note that $M_{i1}(\mathbf{Z}) = \mu_i(\mathbf{Z})$, $M_{it}(\mathbf{Z}) = M_{it-1}(\mathbf{Z}) + \lambda_{it}(\mathbf{Z})$, where $\mu_i(\mathbf{Z})$ is the *ex-ante* bargaining power of person i when the marriage is formed. Finally, it is important to highlight the fact that the eligibility to the cash transfers is one of such distribution factors, i.e. $\mathbf{Z} = (z^*, \mathbf{Z}_{-1})$, where $\mathbf{Z}_{-1}$ is the vector of distribution factors without the first element.

Three points are worth discussion. First, in a dynamic collective household context, distribution factors are variables that affect the allocation of resources by changing the relative weights of the two parents within the household through a change of the outside option. They play a key role in our analysis because they can be used both to test commitment and efficiency. Second, under the assumption of *ex ante* efficiency, only the decision power at the time of household formation $\mu_i(\mathbf{Z})$ may affect household behavior. The main implication is that, in this case, the set $\mathbf{Z}$ can only include variables known at the time the household is formed. Hence, the policy designed to modify the decision power of household members after the household is formed should have no effect on household decisions. Third, the assumption of *ex-ante* efficiency requires that household members

conducted on poor marginalized rural households in Mexico and the data at hand are in the form of a short panel.

can commit at period 1 to an allocation of resources for each future period. Given that this assumption may be too restrictive, the LIC model assumes that intertemporal decisions are a function of individual power at any point in time, not only at the time of marriage. This assumption is explicited in the model through the inclusion of the participation constraint $\lambda_{it}(\mathbf{Z})$ and the reservation utility $\bar{u}_{it}(\mathbf{Z})$.

To provide some further insights into how the participation constraint and reservation utility interact in the LIC model, consider the following example. Suppose that in the first period household members get married and start a family. Based on the initial decision power $\mu_i(\mathbf{Z})$, the household determines the optimal allocation of resources for each future period. This is represented by point A on the Pareto frontier of Figure 1 in the Appendix. In the following period, husband and wife consume their income according to this allocation plan until some exogenous event occurs that makes this resource allocation suboptimal for, say, member m. This could be caused by a new policy that assigns a large cash to the mother. This means that, before the policy, the value of her outside option $\bar{u}_{mt}(\mathbf{Z})$ is lower than her current utility in the marriage and hence the participation constraint is not binding, $\lambda_{mt}(\mathbf{Z}) = 0$. However, the new (unexpected) policy has the effect to increase the outside option of m to the point in which "disagreement within-marriage" is now preferred and the participation constraint becomes binding, $\lambda_{it}(\mathbf{Z}) > 0$. However, given that there is still a set of points on the Pareto frontier that provides each spouse with at least their outside option utility, the binding of the participation constraint triggers a renegotiation as soon as the disagreement becomes optimal. The allocation is renegotiated to make spouse m indifferent between the outside option and staying in the household. This goal is achieved by increasing her decision power $M_{mt}(\mathbf{Z})$. From this point onward, the household divides income according to a new plan until some further exogenous shock triggers a new renegotiation. The new plan corresponds to point B in Figure 1. This simple example tells us that consumption decisions at each point in time depend on the individual decision power prevailing in that period and on all the variables having an effect on it.

The solution of program (2) can be written in the form of state-contingent value functions $V_2(\Phi, z_2^*)$, where Φ is the vector of state variables known at the end of period 2: $\Phi = \{\mathbf{p}_t, \mathbf{P}_t\}$. Before writing the regression functions for period 2 consumption allocation based on the LIC model, it is convenient to re-interpret the program in a three-stage form following Chiappori and Mazzocco (2015). This helps us understanding what is the stage of the household behavior that we are going to test in the empirical section.

2.2 Three-stage formulation and empirical predictions

Chiappori and Mazzocco (2015) provide a formulation of household decisions that has the same solution of the LIC model presented above. We use their alternative formulation to better rationalize the empirical tests that will follow. Household behavior can be captured by three different stages. In the first stage, the household makes allocation decisions of its lifetime resources across all periods. This is called *intertemporal stage* because it focuses on the dynamic nature of household decisions. The second stage considers a particular period where, conditional on the amount of resources allocated in the first stage, the household makes optimal decisions on household-level private and public consumption. This is called the *resource allocation stage*. The third and final stage considers the same period as before and, conditional on the decisions made in stage two, it derives the optimal allocation of private goods between spouses. This is denoted as the *intra-household allocation stage*. In what follows we state formally the optimization problem faced by the household in each of these stages in reversed order, and derive the testable empirical predictions that we bring to the data. For ease of the notation, we drop the time subscript (except for stage one) and the state of nature subscript in all stages, as they are implicit in our presentation.

2.2.1 Stage three: Intra-household allocation

In stage three, the *intra-household allocation stage*, a household takes as given an arbitrary amount of household-level private goods $\bar{\mathbf{q}}$ and public goods $\bar{\mathbf{Q}}$ which are optimally chosen in stage two. Then the household chooses the allocation of private goods between spouses by solving the following static problem for each period t:

$$\begin{aligned} U(\bar{Q}, M_i(\mathbf{Z})) = & \max_{\{\mathbf{q}_m, \mathbf{q}_f\}} \quad \sum_i M_i(\mathbf{Z}) u_i(\mathbf{q}_i, \bar{\mathbf{Q}}) \\ \text{s. to} \quad & \mathbf{q}_m + \mathbf{q}_f = \bar{\mathbf{q}} \end{aligned} \tag{3}$$

where $M_i(\mathbf{Z})$ is the individual decision power of the two spouses. The solution to (3) provides information on the welfare generated by the household in period t. The resulting demand equation for a generic good k takes the following form:

$$\theta_k^3 = \xi_k^3(\bar{\mathbf{q}}, M_j(\mathbf{Z}); \mathbf{d}, \epsilon) \tag{4}$$

where the superscript indicates that this is the demand function derived in stage three, whereas $\mathbf{d}$ and ϵ are a set of observable and unobservable characteristics of the household. The crucial aspect of this demand function is the presence of the Pareto weight function $M_i(\mathbf{Z})$ and its functional dependence on distribution factors. Indeed the manner in which $\mathbf{Z}$ affect demand (4) can be used to test Pareto efficiency, the main underlying assumption of collective models.

Bourguignon et al. (2009) derive necessary and sufficient conditions for collective rationality which are powerful because they are valid for any type of good, either private or public. In order to understand the theoretical restriction that we want to test, we introduce the concept of *z-conditional demand* function. Consider the demand for good l resulting from program (3), $\theta_l^3 = \xi_l^3(\bar{\mathbf{q}}, \mathbf{Z})$, where some of the elements of $\mathbf{Z}$ may not be observed but at least one is. In particular, assume that there is at least one good l and one observable distribution factor z_1 such that $\xi_l^3(\bar{\mathbf{q}}, \mathbf{Z})$ is strictly monotone in z_1. Given strict monotonicity, the demand function for good l can be inverted on this factor: $z_1 = \zeta(\bar{\mathbf{q}}, \mathbf{Z}_{-1}, \theta_l^3)$. We can now define the following:

Definition 1. *The demand function for any good k, private or public, in stage three, is a z-conditional demand if:*

$$\theta_k^3 = \xi_k^3(\bar{\boldsymbol{q}}, z_1, \boldsymbol{Z}_{-1}) = \xi_k^3[\bar{\boldsymbol{q}}, \zeta(\bar{\boldsymbol{q}}, \boldsymbol{Z}_{-1}), \theta_l^3, \boldsymbol{Z}_{-1}] = \varphi_k^3(\bar{\boldsymbol{q}}, \boldsymbol{Z}_{-1}, \theta_l^3) \tag{5}$$

In other words, the demand for good k can be written as a function of total expenditure $\bar{\mathbf{q}}$, all distribution factors but the first, $\mathbf{Z}_{-1}$, and the quantity demanded for good l. Although conditional demands are often used in demand analysis, it is useful to refer to it as z-demands because it incorporates the idea that distribution factors play a central role in the intra-household allocation stage of collective models. Empirically, the restriction that involves the z-conditional demand says that if there exists a distribution factor such that:

$$\frac{\partial \theta_k^3}{\partial z_1} \neq 0 \quad \forall k \tag{6}$$

the demand for good k is compatible with collective rationality if and only if there exists at least one good l such that:

$$\frac{\partial \varphi_k^3(\bar{\mathbf{q}}, \mathbf{Z}_{-1}, \theta_l^3)}{\partial z_p} = 0 \quad \forall k \neq l \quad \text{and} \quad p = 2, \ldots, s \tag{7}$$

The meaning of this testable restriction is the following. If we invert the demand for good l on a relevant distribution factor z_1, which is relevant also for any other good $k \neq l$, and we replace this demand into the demand of any other good $k \neq l$, the effect of any second distribution factor z_p is going to be irrelevant. The intuition is that, by definition, distribution factors affect demand only through their effect upon the location of the final outcome on the Pareto frontier. They do not shift the Pareto frontier per se. Importantly, the effect of the bargaining weight is one-dimensional. Once the location on the Pareto set has been changed by the effect of one distribution factor, the information brought by any other additional distribution factor is uninformative[8].

Dauphin et al. (2015) generalize the approach by Bourguignon et al. (2009) by relaxing one of their critical assumptions, namely that at least one distribution factor

[8]Note that Proposition 2 in Bourguignon et al. (2009) provides three equivalent conditions necessary and sufficient for collective rationality. Empirically, the restriction that involves the z-conditional demand is the most powerful because we can employ single equation methods which are more robust than tests of the equality of parameters across equations. This is the reason why in the present paper we emply this restriction.

(locally) affects all the demand functions; that is, equation (6) in the above notation. They derive the appropriate test procedure starting from the assumption that distribution factors need not locally influence more than two demand functions to yield falsifiable restrictions. More precisely, a system of 2 or more demand functions and 2 or more distribution factors, of a household with 2 decision makers, is locally compatible with collective rationality if:

$$\frac{\partial \xi_k^3(\bar{\mathbf{q}}, M_j(\mathbf{Z}); \mathbf{d}, \epsilon)}{\partial z_p} = 0 \quad \text{or} \tag{8a}$$

$$\frac{\partial \xi_k^3(\bar{\mathbf{q}}, M_j(\mathbf{Z}); \mathbf{d}, \epsilon)}{\partial z_p} \neq\neq 0 \quad \forall k = 1, \ldots, n \quad \text{and} \quad \forall p = 1, \ldots, s \tag{8b}$$

In other words, a demand system (4) that responds to at least two distribution factors is compatible with collective rationality if each of its demands either does not respond to any distribution factors (8a) or responds to all of the distribution factors (8b). Notice that if all the demand functions (4) do not respond to any distribution factors, the data at hand are compatible with unitary rationality.

2.2.2 Stage two: Resource allocation

In stage two, the *resource allocation stage*, the household enters this period with an arbitrary amount of resources x and uses the indirect utility function derived in stage three $U(\bar{\mathbf{Q}}, M_i(\mathbf{Z}))$ to determine the optimal choice of the household-level private and public consumption. Formally, the household solves the following static problem:

$$\begin{aligned} V(x, \mathbf{p}, \mathbf{P}, M_i(\mathbf{Z})) &= \max_{\{\mathbf{q},\mathbf{Q}\}} \; U(\mathbf{q}, \mathbf{Q}, M_i(\mathbf{Z})) \\ \text{s. to} \;\; \mathbf{p}'\mathbf{q} + \mathbf{P}'\mathbf{Q} &= x \end{aligned} \tag{9}$$

where $V(x, \mathbf{p}, \mathbf{P}, M_i(\mathbf{Z}))$ is the indirect utility function for this stage, measuring the household welfare if x resources are allocated to this period.

Under the assumption that each decision maker's power $M_i(\mathbf{Z})$ is fixed, we can reinterpret program (9) as a problem faced by a unique entity that maximizes a weighted sum of individual preferences. Alternatively, we can assume that the decision power is in the hands of one individual (e.g. $M_m(\mathbf{Z}) = 1$ and $M_f(\mathbf{Z}) = 0$) and this benevolent dictator makes optimal decisions for all members of the household. Both restrictions are implied by the so called unitary model of the household (e.g. Becker, 1991) and generate the "income-pooling property" according to which, after controlling for household income, any distributional variable should have no effect on household decisions. This restriction allows us to derive a demand function for stage two of the form $\theta_k^2 = \xi_k^2(x, \mathbf{Z}; \mathbf{d}, \epsilon)$, where the testable formal restriction implied by the unitary model is that:

$$\left.\frac{\partial \theta_k^2}{\partial z_p}\right|_x = 0 \quad \forall z_p = 1, \ldots, s \tag{10}$$

where z_p indicates any distribution factor realized at the time of marriage or afterward. This testable restriction tells us that, at each point in time and with a cross-sectional dataset, static collective models can be used to study the effect of differences in decision power across households on consumption choices. If we reject the null hypothesis implied by (10), we reject the "income-pooling property" and hence the unitary model. This is the most popular test that has been run in the literature of unitary versus collective models (e.g. Attanasio and Lechene, 2002).

2.2.3 Stage one: Intertemporal allocation

In stage one, the *intertemporal stage*, the household deals with the allocation of lifetime resources using the indirect utility function derived in the *resource allocation stage*. The household considers the following problem:

$$\begin{aligned} \max_{\{y_t\}} \quad & \sum_t E_0[V(x_t, \mathbf{p}_t, \mathbf{P}_t, M_{it}(z_2^*, \mathbf{Z}_{-1})) - \sum_i \lambda_{it}(z_2^*, \mathbf{Z}_{-1}) \cdot \bar{u}_{it}(\mathbf{Z}_{-1}, z_2^*)] \\ \text{s. to} \quad & \sum_t x_t + z_2^* = X \end{aligned} \tag{11}$$

where the key variable in the model is the distribution factor z_2^*, which may affect the *ex-post* reservation utility of member m after installation of the policy.

The LIC model is based on the assumption that households cannot commit to the same resource allocation plan made at the time of marriage, irrespective of the events that may occur in the future. Vice versa, under full commitment, only the intra-household decision power at the time of household formation can affect family behavior, and changes in the outside options following marriage have no consequences. This assumption allows to distinguish the LIC model from its special case named Full-commitment Intertemporal (FIC) model (see Mazzocco, 2007).

Given our set-up, we can test for commitment in the following way. Commitment means in practice that the Pareto weight in (11) are constant over time, that is: $M_{it}(z_2^*, \mathbf{Z}_{-1}) = \mu_i(\mathbf{Z}_{-1}) \ \forall t$, where $\mu_i(\mathbf{Z}_{-1})$ is the *ex-ante* bargaining power of person i when the marriage is formed. With this assumption, and 2-period horizon, we can solve the model recursively, and the resulting demand equation for stage one of a generic good k at period 2 takes the following form:

$$\theta_k^1 = \xi_k^1(x, \mu_i(\mathbf{Z}_{-1}), z_2^*; \mathbf{d}, \epsilon) \tag{12}$$

The key empirical implication of commitment in our context derives from the fact that the policy realizations z_2^* could not be forecast at the time of marriage and hence it enters the problem only through the period-2 budget constraint. In the no-commitment case, being eligible to the policy may have two effects: a wealth effect, as in the commitment case, and a change in the reservation utility for m that leads to a shift in bargaining power. Hence, if we assume full commitment, after controlling for the wealth effect, there should be no remaining effect in the demand of goods. Technically, the testable restriction is:

$$\left.\frac{\partial \theta_k^1}{\partial z_2^*}\right|_{\tilde{x}} = 0 \tag{13}$$

where $\tilde{x}$ indicates the optimal amount of income allocated for period 2 by the household augmented with the unexpected cash transfers entitled to the mother. The difference between (10) and (13) is that in the latter we are using exclusively a distribution factor that has affected the family in the second period, after the marriage was formed. Note that a rejection of the null hypothesis in (13) corresponds to an automatic rejection of the null hypothesis in (10).

3 Data

We investigate how households respond to monetary incentives using a sample drawn from the surveys collected to evaluate the impact of PROGRESA. This is a conditional cash transfers (CCT) program implemented in rural Mexico in the late 1990s. The choice of this dataset is motivated by a variety of reasons. First, the monetary incentives were quite large and had a real bite on households' behavior inducing them to change their consumption patterns. Second, the surveys are very detailed and of high quality allowing us to construct vectors of quantity and prices for various important commodities. Third, the available dataset contains three distribution factors which, altogether, allow us to test the main hypothesis outlined in the theory part. The remaining of this section is divided in two subsections. First, we provide some background information on the program. Second, we present the evaluation surveys, how prices and quantities are aggregated, and some descriptive statistics of the sample adopted in our empirical analysis.

3.1 Program design

The original PROGRESA program was implemented between April 1998 and December 2000. Later it was extended to include new households both in rural and urban areas. From its start, PROGRESA was subject of a rigorous evaluation based on random assignment. 10.000 localities were included in the first expansion phase, and from here 506 were selected in the evaluation sample, 320 of them were randomly chosen to have an early start of the program, whereas the remaining 186 formed the control group. In practice, households in the these latter villages were not included in the program until late 1999, which means that they became eligible for the grant only after this date. "Eligible" households in treatment villages started receiving the cash transfers subject to the appropriate conditionality already in April 1998. Whereas "eligible" households in control villages were still observed during this time but they started benefiting from the payment (in the same manner) after November 1999.

The main objectives of the program were to introduce incentives to improve the accumulation of human capital of children and at the same time to alleviate short-term poverty. To be eligible, a household must be sufficiently poor (in the program sense). The transfers were paid to the mother every two months and were largely in the form of scholarships to four grades of primary school except the first two and the initial three grades of secondary school. These transfers are conditional on certain behaviors: first, children must attend at least 85% of classes; second, household members must undergoing periodical health checks; third, the transfer recipients must attend nutrition and health classes. The strong involvement of the mother in the program was motivated by the assumption that they have a stronger taste for child well-being and are more responsible at managing households resources. Moreover, a change in relative income of spouses was motivated by the desire to change the position of woman within rural families in Mexico, which was the intended by-product of the intervention.

The program take-up rate was high. More than 95% of households complied to the program and contamination of the control group was negligible. The program was later expanded to other households in rural areas who were followed throughout the 2000s. The program was also expanded to urban areas and changed its name into Oportunidades, although the basic set up of the original sample remained. The implementation of PROGRESA is considered a success for many reasons and this led other countries, both in Latin America, Asia, Africa, and some developed countries as well, to adopt a similar approach that made it very popular. When PROGRESA was designed in 1997, 300.000 families were benefiting from it. Today Oportunidades covers over 25 million individuals (5 millions households) representing 25% of the Mexican population. This welfare program represents today the policy with the largest budget allocated in this country. PROGRESA has been found to increase education attainment (Schultz, 2004; Attanasio et al., 2012), to decrease short term poverty (Tommasi and Wolf, 2016), and to improve health (Gertler, 2004; Behrman and Parker, 2011). Detailed information on the program and its evaluations can be found in Skoufias et al. (2001), Skoufias (2005) and Fiszbein et al. (2009).

3.2 Sample selection and descriptive statistics

During the early expansion phase of PROGRESA, the government collected two surveys prior to intervention, in 1997 and March 1998, and four surveys post interventions, in October 1998, May 1999, November 1999, and November 2000. The first two surveys cannot be used for the purpose of our exercise, as they do not contain information on prices and quantities of the commodities purchased. In the present paper we use two of such surveys, October 1998 and May 1999, which were collected 6 months and 12 months, respectively, after the households started receiving the cash transfers. The surveys included detailed information on expenditures at the household level and detailed information on members of the household. The original evaluation sample contains 24.077 households of which 61.5% are natural parents with any number of children and no other adult individual living in the household.

In order to have an homogeneous sample on which we can test the hypothesis of interest, we select a sub-sample that satisfies the following restrictions. First, there are only households with both natural parents in our sample and one to at most six children. This means that households with at least one other adult member are excluded and the mother is always the recipient of the cash transfers. Second, households with children aged 17 or above are also excluded from the sample, in order to avoid to have multiple decision makers besides the parents. The resulting sample consists of 5.125 households observed in 1998 and 4.932 households observed in 1999. In the Online Appendix we present the means of various baseline household-level characteristics for eligible households in treatment and control villages. For a formal comparison of the averages of the two groups to show balance see Behrman and Todd (1999). As we can see from the table in the appendix, households are disadvantaged in a number of important ways. First, the education of head and spouse is quite low, as the average adult has slightly more than a primary school diploma[9]. Second, families are quite large as the average number of children is 4. Third, 37% of households have indigenous origin. Finally, only about a quarter of localities have a secondary school in the village.

We are interested in studying the household responses to cash transfers in terms of demand for food components, which, in our sample, represents about 80% of non-durable expenditure. The PROGRESA data contains very detailed information on both expenditure and consumption for many (narrowly defined) commodities. Following Attanasio and Lechene (2014), we use aggregated data to create budget shares of 5 different commodities: starches; pulses; fruit and vegetables; meat, fish and diary; and other foods. As explained in details by the authors, for each of the individual commodities that make the 5 commodities that we use, consumption is computed as to include both what has been bought and quantities obtained from own production, payments in kind and gifts.

[9]Education of head and spouse is coded as follows: 1 for incomplete primary, 2 for primary, 3 for incomplete secondary, 4 for secondary and above.

The quantities are valued in pesos using locality level price information derived from unit values. As reported in both papers, the authors take particular care to avoid duplication induced by household production. For instance, if a household has consumed a good that was produced at home, they include the value of this good (valued at average prices in the town) but do not include the value of the raw material that was purchased to make that good.

In order to estimate the demand system we need to compute unit values of the five commodities. These are used to evaluate consumption in kind and to compute price indexes for each of the composite commodities. Unit values are computed for each household dividing the value of the purchase by its quantity. The value of the purchase commodity is computed by using village-level prices for individual commodities, where the village-level price is selected by looking at median unit value of the households that purchased that product in a given locality. More details on the computation of these unit values and how price indexes are constructed can be found in Attanasio et al. (2009). To get a precise parameter estimates of the demand system, it is notoriously required to have a sufficiently large variation in prices of the commodities of interest. As it is well known from the applications of the PROGRESA, there is considerable variation in prices across villages and time in the data. Notice that although prices of foods decreased between October 1998 and May 1999, this trend is not different between treatment and control villages.

4 Empirical implementation

Our aim is to test how households allocate their available resources after receiving a large exogenous cash transfers. Household budget shares of food commodities are the main output of interest. The demand for it is modeled assuming separability of these goods from the non-food consumption and labor supply. We focus on demand for food for a variety of reasons. First and for most, food consumption is the most important commodity in the budget as it accounts for around 80% of the expenditure of the households in the sample. Second, prices for the non-food consumption are not observed and hence it is practically not possible to use these goods. However, exactly because food shares is so important, a test of the collective model on these commodities remains a meaningful exercise. In what follows, the first sub-section discusses the appropriateness of the three distribution factors that we use to test commitment and efficiency of the resource allocation. These are the most important variables for the purpose of our exercise. The second sub-section discusses the consumption behavior of our sample, that is, household preferences and the observed demand equations. The third sub-section deals with the estimation strategy and the methodological issues that have been raised in the literature when one aims to identify the relationship of interest with data coming from a cash transfers program such as PROGRESA (e.g. Attanasio and Lechene, 2002; 2014; Attanasio et al. 2012, 2013). In our context, we are particularly concerned with the endogeneity of both total expenditure and the number of children enrolled in secondary school.

4.1 Distribution factors

In order to test the hypothesis of interest, we need to find at least two variables that we assume to affect the allocation of resources but not preferences. These variables are called distribution factors and enter the Pareto weight function of the two agents within the household. Browning et al. (2014) report the most common distribution factors used in the literature. As the authors argue, it is a difficult exercise to find plausible distribution factors because theory does not give guidance as to what constitutes a distribution factor and for each variable it is possible to find a reason why it could affect also preferences or the budget constraint. In order to conduct a robust analysis, in the present paper we use three, and not two, of the most credible distribution factors used in the literature. In what follows we describe and motivate each of these variables.

The first distribution factor used is the eligibility to PROGRESA within a village. This is a dummy variable taking value 1 if the household belongs to a treated village and 0 otherwise. Since the grant is targeted to the mother, receipt of the transfers constitutes an exogenous increase in the share of the household income that she controls. This share of income is not an argument of preferences, and conditional on total resources available, it does not affect the budget constraint. Given the random assignment of the program, the treatment variable constitutes an ideal distribution factor, which explains why it has been used so often in the recent literature to test the collective model. Two remarks on the treatment dummy are in order. First, the grant affects not only the distribution of resources within the household but also the total resources available. This implies that we need an appropriate specification of the demand system to control for total resources available after the treatment. Conditional on all the resources, including also those coming from the program, the receipt of PROGRESA shoud make no difference to the allocation of household resources among different commodities. If instead, after conditioning, the grant has a residual effect on allocation, it must be because it has shifted the Engel curves as a consequence of a shift in Pareto weights. Second, the PROGRESA grant is a conditional cash transfer, where the most stringent conditionality is the child school enrollment. In the case of the Mexican context, the conditionality is not stringent for families who have to enroll their children to primary school, as primary school enrollment is almost universal. It is however stringent in the case of families with children going to secondary school. A correct specification of the demand system needs to account for this as well[10].

[10]In principle, there is a third remark that should be made. If the PROGRESA grant affected labor

The second distribution factor used is the relative importance of household's and wife's family network in the village. This information was collected by Angelucci et al. (2009) and used as a distribution factor to test the collective model by Attanasio and Lechene (2014). The main idea behind the use of the network information is the fact that a stronger presence of family members in the village affects the individual value of their outside option. Indeed, as the authors argue, it is possible that the relative weights of husband and wife in the allocation of resources depend, within the context of poor marginalized rural households, on the relative strength and influence of the two extended families in the village. The relative importance of the spouse's networks is constructed by Angelucci et al. (2009) as follows. The authors exploit the fact that the PROGRESA evaluation surveys are a census of each village and the convention of Spanish last names to map the network of relatives within each community. Indeed, in Spanish-speaking countries, individuals get two surnames, the first one from the father and the second one from the mother. Using the PROGRESA surveys it is possible to know the number of relatives, for each adult, that are present in the village. The relative importance of husband and wife's networks is then constructed in two ways: the size and wealth of the networks[11]. For a further discussion of the appropriateness of this variable as distribution factor, see Section 4.2 of Attanasio and Lechene (2014).

In practice, there is no need to find another distribution factor, as both the z-conditional test and the "all or nothing" restriction test require only two of such variables. However, part of our analysis is to study the heterogeneity of the efficiency hypothesis, in particular separately between the treatment and control groups. In this case it is not possible to use the treatment indicator as a feasible distribution factor. Angelucci and Garlik (2015) use sex ratio in the village as the conditioning distribution factor. Since in our dataset we do not have such information, we proceed in a different way and use education difference between spouses. As explained at length by Quisumbing and Maluccio (2003), conditional on the education *level* of the spouses, the *difference* is an attractive measure of bargaining power because it captures some cultural aspects within a family. This is one of the main reason why it is commonly used in the literature (e.g. Gitter and Barham, 2008; Schady and Rosero, 2010; Tommasi, 2015). Human capital differences are sound distribution factors because they are predetermined and do not change after the introduction of the grant. Moreover, it is an endogenous variable in a model of marriage market selection, because it is potentially correlated with other observables that guides the selection of the partner. However, it is clearly exogenous to decisions made *within*

supply, then it would not be a valid distribution factor. However, it has been shown that the grant did not have any effect on labor supply of adult members and hence it is likely that this does not constitute a problem in our specific case (Skoufias, 2001; Skoufias and Di Maro, 2008).

[11]More formally, for each individual $i = m, f$, they construct the relative size of the networks as the ratio of $n_i/n_m + n_f$, where n_i is either the number of relatives in the village or the value of their wealth for each individual i. Wealth is proxied by (food) consumption of individual's relatives.

marriage (Quisumbing and Maluccio, 2003). Notice that there is a further reason why it is useful to use a third distribution factor. Some of the results in the literature testing collective rationality have been criticized on the basis of which distribution factors were employed in the analysis. Hence having more than two of such variables allows us to obtain more robust results, as we can evaluate the two tests of efficiency by looking at any pairwise combination of the two variables and verify whether we obtain consistent results.

4.2 Functional forms

In our empirical application we assume that households have preferences given by the Integrable QAIDS demand system of Banks et al. (1997). QAIDS is very convenient because it allows flexible prices responses, the quadratic income allows the Engel curves to display a great variety of shapes and at the same time the system of demand equations derived preserves theoretical consistency. Moreover, in the context of PROGRESA data it has been shown by several authors that it is important to allow for quadratic responses of food budget shares (Attanasio et al., 2009; 2013).

The indirect utility function of each household is assumed to be of the following form:

$$v = \left\{ \left[\frac{\ln x - \ln a(\mathbf{p})}{b(\mathbf{p})} \right]^{-1} + \lambda(\mathbf{p}) \right\}^{-1} \tag{14}$$

where

$$\begin{aligned} \ln a'(\mathbf{p}) &= \alpha_0 + \sum_{k=1}^{n} \alpha_k \ln p_k + \frac{1}{2} \sum_{k=1}^{n} \sum_{j=1}^{n} \gamma_{kj} \ln p_k \ln p_j \\ b(\mathbf{p}) &= \prod_{k=1}^{n} p_k^{\beta_k} \\ \lambda(\mathbf{p}) &= \sum_{k=1}^{n} \lambda_k \ln p_k \end{aligned} \tag{15}$$

The parameters α_k, β_k, λ_k and γ_{kj} $(\forall j, k)$ are to be estimated. Adding up requires that $\sum_k \alpha_k = 1$, $\sum_k \beta_k = 0$, $\sum_k \lambda_k = 0$ and $\sum_k \gamma_{kj} = 0$ $(\forall j)$. Homogeneity is satisfied if $\sum_j \gamma_{kj} = 0$ $(\forall k)$. Slutsky symmetry is satisfied if $\gamma_{kj} = \gamma_{jk}$ $(\forall j, k)$[12]. Notice that the indirect utility function underlying Deaton and Muellbaur's (1980) Almost Ideal Demand System corresponds to equation (14) where $\lambda_k = 0$ for all goods. Applying Roy's identify

[12]Attanasio and Lechene (2014) estimate their QAIDS without imposing the Slutsky symmetry condition since in a collective model of the household it does not hold (Browning and Chiappori, 1998). Whereas in Attanasio et al. (2009, 2013) the authors use the same dataset to estimate the demand for food and they do impose the Slutsky symmetry. Not imposing it goes already towards the direction of having in mind a collective model of the household. However, since in our results we do reject the collective model, we decide to imposing it. Notice that not imposing it does not change the qualitative results anyway.

to equation (14) we obtain the QAIDS budget share equations for each household and commodity k ($k = 1, \ldots, n$):

$$w_k = \frac{\theta_k^s}{x} = \alpha_0 + \phi'\mathbf{d} + \psi'\mathbf{Z} + \sum_{j=1}^{k} \gamma_{ij} \ln p_j + \beta_k \ln \left\{ \frac{x}{a(\mathbf{p})} \right\} + \frac{\lambda_k}{b(\mathbf{p})} \left[\ln \left\{ \frac{x}{a(\mathbf{p})} \right\} \right]^2 + \epsilon_k \quad (16)$$

where w_k indicates the kth budget share of a household facing a price vector $\mathbf{p}$ and total expenditure level x, whereas $\mathbf{d}$ and $\mathbf{Z}$ are vectors of, respectively, individual demographic characteristics and distribution factors[13]. In order to link equation (16) with the theoretical section, notice that each household budget share is given by the demand for food θ_k^s divided by total expenditure x. Here the superscript s indicates one of the three stages of the household decision process in the LIC model, depending on which stage of the household allocation process we are looking at.

4.3 z-conditional demand system

In order to estimate the z-conditional demand system (5) for good θ_k^3, we have to allow that the conditioning good θ_l^3 might be endogenous. This problem can be overcome because the excluded distribution factor on which the demand is inverted becomes a natural instrument for θ_l^3. Let N, the relative family network, be the excluded distribution factor. The demand function for commodity k ($k = 1, \ldots, n$), in stage three, of the household allocation process can be inverted on this factor:

$$N = \frac{1}{\psi_N}\theta_l^3 - \frac{\alpha'}{\psi_N}\mathbf{Z}_{-1} - \frac{\beta}{\psi_N}f(x) - \frac{\mu'}{\psi_N}\mathbf{d} - \frac{1}{\psi_N}u_l$$

where now $\mathbf{Z}_{-1}$ contains only the two remaining distribution factors. Substituting this equation for N in the demand for all other goods results in the system of z-conditional demand functions:

$$\theta_k^3 = \tilde{\alpha}'\mathbf{Z}_{-1} + \tilde{\gamma}\theta_l^3 + \tilde{\beta}f(x) + \tilde{\mu}'\mathbf{d} + \tilde{u}_k \quad (17)$$

for all goods $k \neq l$. The test of collective rationality becomes a test of the significance of $\tilde{\alpha}$.

The choice of the conditioning distribution factor and conditioning good is crucial for the reliability of the empirical results. Theory tells us that the conditioning distribution factors must be relevant and must affect the conditioning good monotonically. In

[13]The impact of these variables runs through the coefficients ϕ and ψ, whose estimates constitutes the main purpose of our empirical investigation. In principle both vectors $\mathbf{d}$ and $\mathbf{Z}$ could affect the demand system in other ways, not necessarily through the intercept only. As a robustness check, we re-estimated the parameters of a general QAIDS model where demographic characteristics and distribution factors were allowed to change the curvature of the demand system in multiple ways. Almost all the additional parameters were not significant, which indicates that it is not reductive to focus only on changes in the intercept. Results are available upon request.

the empirical analysis we use the network variable as our preferred conditioning distribution factor which has been shown by Attanasio and Lechene (2014) to satisfy all the requirements by the theory. As a robustness check we use also our variable of education difference between spouses. As it will be clear from Section 5, the qualitative results of the analysis do not change irrespective of the choice of the conditioning distribution factor.

4.4 Handling endogeneity

Since the dataset used comes from the evaluation of a cash transfers program, which has some important conditionality attached, the main methodological concern in estimating the demand system (16) is the endogeneity of total expenditure and child school enrollment. A further methodological concern is the non-linearity of the system which makes the recovery of the parameter estimates more complicated. The latter issue is tackled by estimating the complete system with the iterated Feasible Generalized Non-Linear Least Squares (FGNLS) estimator. The former concern is tackled with a control function approach as it is commonly applied in demand analysis (e.g. Blundell and Robin, 1999). In the following paragraphs we explain the concern for each of the endogenous variables and how we deal with it.

For the endogeneity of total expenditure, notice that the implicit assumption behind our exercise is the idea that households decide their budget structure under two-stage budgeting: first they decide how much to allocate to food and then how much to allocate to each of the 5 components of food. The residuals in (16) can be interpreted as household's unobserved tastes that affect each budget share. There are two main arguments in the literature for why total expenditure x should be endogenous. One is that taste shocks that determine total expenditure x may be correlated to the unobserved shocks to a particular food component in the system. The other one is that measurement errors in the budget shares may be correlated with measurement error of total expenditure. In the present paper we follow Attanasio and Lechene (2002, 2014) and instrument total expenditure x with the average agricultural wage in a village. This is a strong instrument and the implicit assumption in using it is that any measurement error in village-level income is not correlated with measurement error of household total expenditure, which is a commonly accepted assumption. As the authors explain at lenght, this is a valid instrument if labor supply is separable from consumption. With respect to this, there is a large evidence that PROGRESA did not affect adult labor supply and hence it is not of a concern for us (Skoufias, 2001; Skoufias and Di Maro, 2008).

The second endogenous variable in system (16) is the number of children enrolled in school. As it was explained before, eligible households receive a (large) portion of the grant if their children are enrolled and attend school. This conditionality requirements, which is controlled in the demand equations, might affect consumption behavior if sending children to school imposes additional costs like books, uniforms, etc. Moreover, if children are fed in school, this would further impact the budget share of food. It is plausible to believe that the number of children in school is endogenous in the system because the unobserved taste for school may be correlated with unobserved taste for certain foods. Notice that the concern is only for the number of children in secondary school as the enrollment in primary school is almost universal in rural Mexico and hence not affected by the grant. In order to allow for endogeneity of children in secondary school, we follow Attanasio and Lechene (2002, 2014) and instrument it with a dummy variable indicating the existence of a secondary school in the village and with the distance from the closest secondary school if this is not present in the village. The implicit assumption made is that these two instrumental variables affect the schooling decisions of parents but not directly the structure of their expenditure on food.

Finally, before concluding this section, it is worth noticing that the QAIDS budget share equations of the z-conditional demand depicted in equation (17) contains a third endogenous variable: the budget share of the conditioning good. As the conditioning good θ_l^3 is correlated with the unobserved taste shock of the demand for good θ_k^3, this needs to be instrumented for. The natural instrument to use is already suggested by the theory and by the z-conditional demand test that we run: the distribution factor used to invert the demand of the conditioning good satisfies the common requirements for valid instrumental variables. Hence, in estimating equation (17) we apply the same control function approach as before adding the residuals from the first stage of the conditioning good as well.

5 Results

We divide this section in three parts. First, we present the results of the test of commitment. Second, we present the results of the test of efficiency, which represents our core section, and run several robustness checks to further validate our empirical results. Finally, we discuss the implications of our results and reconcile them with the existing literature. In all specifications we instrument total food expenditure with village-level agricultural wage (and its square), and number of children in secondary school with a dummy if there is a secondary school present in the village and distance to the closest secondary school. We control for a large set of pre-treatment village, household and individual characteristics. Village characteristics include town size and prices. Household characteristics include number of young children, number of children enrolled in primary school, number of children enrolled in secondary school, number of relatives eating in the household and number of household members eating outside the household. Individual characteristics include the level of education of both parents, age of the household head and an indigenous head dummy. All the standard errors are clustered at village level and bootstrapped 200 times.

5.1 Test of commitment

Section 2 makes it clear that in order to test commitment we require a distribution factor that was unexpected at the time of marriage formation. In our set up this variable is represented by the eligibility to PROGRESA. We estimate system (16) where the corresponding demand equations that give rise to these household's budget shares are depicted in equation (12) of the theoretical model. Moreover, the empirical test that we wish to run corresponds to the restriction implied by equation (13). Table 2 reports the p-values of the joint significance tests of treatment for 6 different specifications of system (16), separately for 1998, 1999, and the pooled sample. The blue boxes indicate when the test yields p-values below the 10% level, hence when we reject the null hypothesis that the joint test is not significant. These specifications differ only for the distinct combinations of distribution factors that we use. We do this in order to study the sensitivity of the test to the interaction of the main variable, treatment, with the other variables that affect the Pareto weights. As it was expected, the treatment indicator is significant in all specifications and in both years. These results clearly show the inconsistency of the predictions of the Full-commitment Intertemporal Collective (FIC) model of household behavior.

Next, Table 3 investigates further the sensitivity of the commitment hypothesis by re-estimating system (16) and re-running restriction (13) for 7 sub-samples, separately for 1998, 1999, and the pooled sample. The different sub-samples are described as follows. The first three are households where either the father is more educated than the mother, or where the spouses have the same level of education or where the father is less educated. The next four sub-samples are households differentiated by the age of the father: younger than 30 years old, between 30 and 40, between 40 and 50 or older than 50. The table reports the p-values of the joint significant test of the distribution factors. Like before the blue boxes indicate when the test yields p-values below the 10% level. As it is clear from the results, the hypothesis of commitment is now rejected only for the households in 1999, not in 1998. Given that these estimates refer to the same population of interest, with the same number of clusters, variability and characteristics, the only possible explanation to reconcile the results of Table 3 with Table 2 is that the effect size of the program in 1998 is smaller than the minimum detectable effect size (MDE) of the randomized experiment in this year[14] (Bloom, 2006). Whereas in 1999, the effect size becomes larger than the MDE and our test procedure allows us to detect it. These results have a clear

[14]We typically define an MDE as the smallest true effect that has 80% power for a two-tailed test of statistical significance at the 0.05 level.

theoretical interpretation in light of the model that we presented in Section 2. We can argue that commitment is very strong in rural Mexico and only a large exogenous shock of the outside option can lead the mother to gain bargaining power which triggers a renegotiation of the resource plan within the family. In the Online Appendix we provide both a statistical explanation for these estimates and a theoretical interpretation for why the Pareto weight takes time to shift.

Two final remarks are in order. First, the test of commitment that we just presented has been interpreted in the literature as a test of the unitary model (Attanasio and Lechene, 2002; Rubalcava et al., 2009). The dynamic setting that motivates the empirical implementation in the present paper makes it clear that, technically, this is a test of commitment rather than of the unitary model. However, as Section 2.2.3 explains, a rejection of the null hypothesis in (13) corresponds to an automatic rejection of the null hypothesis in (10): that is, a rejection of the FIC model implies a rejection of the unitary model. Second, Table 2 and 3 report also the p-values of the joint significance tests of other distribution factors different from treatment. Each of these tests corresponds to the restriction implied by equation (10) of the theoretical section. This is the proper test of the unitary model which has been implemented several times in the literature. As we can see, the model is rejected also with these further variables.

5.2 Test of efficiency

The rejection of the FIC model of the household leads us to the question of whether households re-negotiate efficiently, that is, whether they are ex-post efficient and behave according to the collective model. For both 1998, 1999 and the pooled sample, we test collective rationality using jointly the procedure proposed by Dauphin et al. (2015) and Bourguignon et al. (2009). We use the latter regardless of whether one of the distribution factors locally affects each demand equation of the system, which mimics the common practice in the literature. Concerning the former test, the "all or nothing" restriction is a necessary condition requiring that each demand function either do not respond to any of the distribution factor or to all of the distribution factors, and requires that we have available at least two of these variables. In practical terms, we estimate system (16) and the empirical test that we wish to run corresponds to the restriction implied by equation (8b) for both distribution factors. Concerning the latter test, in practical terms we estimate system (17) where the corresponding demand equations that give rise to these household's budget shares are depicted in equation (5) of the theoretical model. The empirical test that we wish to run corresponds to the restriction implied by equation (7). In all specifications, the conditioning distribution factor is the relative family network between spouses. Results are twofold.

First, according to the results of the "all or nothing" restriction in Table 4, for any combination of two (out of three) distribution factors we reject the collective model in both 1998, 1999 and the pooled sample. That is, there is no combination of two distribution factors where either both are significant or non significant. Only in 1998 we "almost" fail to reject the collective model, but the rejection is clear for both 1999 and the pooled sample. The results of the z-conditional test in Table 5 lead us instead to a different story. To run this test, we take the demand equations where the distribution factors are most responsive. This is required to increase the validity of the test results. As we can see, we have evidence in favor of the collective model in 1998, but not in 1999. And we have some ambiguity against collective rationality in the pooled sample. According to the results of the z-conditional test alone, it would seem that households do make efficient decisions at the beginning of the program, but fail to do so later in time. These results are very much in contrast with respect to the previous literature as discussed at length in the following sub-section. We argue that they are not to be believed as they do not satisfy the necessary condition put forward by Dauphin et al. (2015). The conclusion is that households in the PROGRESA sample *do* waste their resources.

Second, the observed inefficiency in the household resource allocation process may be driven by some specific sub-groups. Given that we have three distribution factors available, we investigate this hypothesis by splitting the sample between treatment and control groups and run the efficiency test separately for both groups using network and education difference as distribution factors. Again the inversion of the demand system is done on the network variable and the efficiency hypothesis is tested using education difference. According to the results of the "all or nothing" restriction for the control group in Table 6, households behave efficiently in 1998, whereas this hypothesis is harder to defend for the 1999 and the pooled sample. The results of the z-conditional test in Table 9 are instead all in support of the efficiency hypothesis for this group. The treatment group, instead, behaves inefficiently according to the "all or nothing" restriction in both 1998, 1999 and pooled sample. These test results are in strong disagreement with the results of the z-conditional test, which favor the efficiency hypothesis. All in all, following Dauphin et al. (2015), there seems to be some indirect evidence, at least for the beginning of the welfare program, that the overall inefficiency is driven by the treatment group. This is not obvious, however, for the 1999 data. The practical implications of these results are quite strong as they seem to suggest that receiving the cash transfers may play some role in the observed inefficient allocation of the resources by the household. With the data at hand it is not possible to investigate further this phenomenon, but these evidence suggest that this non-cooperative behavior of households eligible to receive a large exogenous cash transfers should be further investigated.

Three remarks are needed. First, in Table A2 of the Online Appendix we provide the results of the z-conditional demand test using education difference as conditioning distribution factor instead of network. As we can see, the choice of the conditioning distribution factor does not change the conclusion of our analysis. Second, Angelucci and Garlick (2015) use the Bourguignon et al. (2009)'s test procedure and present evidence of within-sample variation in the efficiency of intra-household resource allocation. In particular, they observe that the consumption patterns of these rural households are Pareto efficient when the heads are relatively old, but not when they are relatively young. We use our approach to investigate whether the results are consistent with their findings and find some similarities. As we can see in the Online Appendix, according to the "all or nothing" restriction test in Table A3 and A4, households with older heads behave efficiency in 1998 but not in 1999, whereas households with younger heads behave inefficiently in both time periods. Notice that, if we ignored the requirement that at least one distribution factor has to be significant across all the equations of the demand system, the results of the z-conditional demand test would tell us that also households with younger heads seem to behave efficiently[15]. Third, the possible weak instruments problem is not of concern in our estimates. The first stage F-test for both total expenditure and enrollment in secondary school, in all specifications and for both years, are always above 20, which clearly satisfies the Kleibergen and Paap (2006) critical values for strength of instruments under heteroschedasticity.

5.3 Reconciling the results with the existing literature

Our results on efficiency are strongly in contrast with respect to the established evidence that the households in the PROGRESA sample respond efficiently to the monetary incentives (Bobonis, 2009; Attanasio and Lechene, 2014; Angelucci and Garlick, 2015). The differences in the results reflect several differences in the implementation of the test which, we argue, may have led the authors to conclusions that are not robust.

First and foremost, all three papers implement the z-conditional demand test by Bourguignon et al. (2009) bypassing the critique of Dauphin et al. (2015). Indeed, none of the paper satisfies the requirement that at least one distribution factor is significant in all equations of the demand system, which leads to weak results. Second, none of the papers implements an actual QAIDS model a-la Banks et al. (1997). Indeed their tests of efficiency are based on the estimate of a linear version of the QAIDS, called $\ell-$QAIDS, which has been shown to be problematic (Matsuda, 2006). Although neither of the efficiency tests requires price variation, or the estimate of the parameters attached to prices, bypassing a proper estimation of the demand system may lead to serious bias in the parameter estimates. This issue is documented in several papers by Alston et al. (1994), Moschini (1995), Asche and Wessells (1997) and Matsuda (2006). The basic issue that these authors point out is the inadequacy of the linearization of the price indices to properly approximate the QAIDS model. As it is further explored in the Online

[15]Results available upon request.

Appendix, linear approximation of the QAIDS requires that $a(\mathbf{p})$ and $b(\mathbf{p})$ be replaced with composite variables, which are free of unknown parameters. The most common composite variable adopted for the approximation is the Stone price index suggested by Deaton and Muellbaur (1980). This is what the authors above have implicitly adopted in the estimate of the QAIDS model before running their efficiency test. However, Moschini (1995) in the context of the AIDS model and Matsuda (2006) in the context of the QAIDS model, show that employing the Stone index instead of the traslog aggregator can seriously bias elasticity estimates partly because this index is influenced by changes in units of measurement.

While these differences are structural, there are also few differences in the sample selection strategy and variables choice. First, similarly to Attanasio and Lechene (2014), but differently from Bobonis (2009) and Angelucci and Garlick (2015), we use only two waves of data after PROGRESA began to distribute cash transfers and focus only on food consumption. The other authors use three waves and model also non-food consumption. The problem with this implementation is that the surveys do not contain information on prices for non-food commodities and hence it is not possible to implement a proper QAIDS model as we just pointed out. Second, again similarly to Attanasio and Lechene (2014), but differently from Bobonis (2009) and Angelucci and Garlick (2015), we use treatment and relative size of husband's and wives family networks as distribution factor, but add relative education differences between spouses as a third distribution factor.

6 Conclusion

The purpose of the present paper is to test some of the main implications of the collective model, which is raising as a more plausible framework to represent the decision making of process of households. The analysis is based on a dataset collected to evaluate the effects of PROGRESA, a CCT program implemented in Rural Mexico in late 90s. This dataset is suited for a variety of reasons. First, the program was targeted to woman in a randomized fashion. Moreover, the monetary incentives were quite large and had a real bite on households' behavior inducing them to change their consumption patters. Finally, the surveys are very detailed and of high quality allowing us to construct vectors of quantity and prices for various important commodities. This set-up, together the richness of the dataset, allows the construction of three valid distribution factors, which are variables that affect demand only through the Pareto weights that define the allocations, and represent the most important variables for our purposes.

We provide new evidence on three issues. First, commitment of spouses to an allocation plan is very strong in rural Mexican households, because it requires a large effect and a high statistical power to be picked up. Second, contrary to the existing literature, we find strong evidence against Pareto efficiency. In other words, the collective model is not suitable to rationalize the observed behavior in the PROGRESA data. Third, we provide some evidence that the observed inefficiency may be driven by the participation into the CCT program. That is, allocating a large amount of money in the hands of the woman may enhance disagreements between the spouses which triggers an inefficient allocation of their resources.

References

[1] Adamowicza, W., Dickie, M., Gerking, S., Veronesi, M. and Zinner, D. (2013). Pareto Efficiency in Intra-household Resource Allocation and Risks to Children's Health. Working paper.

[2] Aguero, J. M. (2008). Are intrahousehold allocations efficient? Evidence from Monte Carlo simulations. Mimeo, University of California, Riverside.

[3] Alston J.M., Foster K.A., Green R.D. (1994). Estimating elasticities with the linear approximate almost ideal demand system: some Monte Carlo results. *Review of Economics and Statistics* 76:351-356.

[4] Angelucci, M. and Garlick, R. (2015). Heterogeneity and aggregation: testing for efficiency in intra-household allocations, mimeo.

[5] Angelucci, M., De Giorgi, G., Rangel, M. and I. Rasul (2009). Village Economies and the Structure of Extended Family Networks, *The B.E. Journal of Economic Analysis and Policy*, vol. 9(1), pages 44.

[6] Apps, P. and Rees, R. (1988). Taxation and the Household. *Journal of Public Economics*, 35.

[7] Asche F. and Wessells, C. (1997). On Price Indices in the Almost Ideal Demand System. *American Journal of Agricultural Economics*, 79 (4): 1182-1185.

[8] Attanasio, O., Battistin, E. and A. Mesnard (2012). Food and Cash Transfers: Evidence from Colombia, *Economic Journal*, Royal Economic Society, vol. 122, pp. 92-124.

[9] Attanasio, O. and Lechene, V. (2002). Test of income pooling in household decisions. *Review of Economic Dynamics* 5(4), 720-748.

[10] Attanasio, O. and Lechene, V. (2014). Efficient responses to targeted cash transfers. *Journal of Political Economy.*

[11] Attanasio, O., Di Maro V., Lechene V and D. Phillips (2013). The effect of increases in food prices on consumption and welfare in rural Mexico, *Journal of Development Economics.*

[12] Aura, S. (2005). Does the balance of power within a family matter? The case of the Retirement Equity Act. *Journal of Public Economics* 89:1699-1717.

[13] Banks, J., Blundell, R. and A. Lewbel, (1997). Quadratic Engel Curves And Consumer Demand, *The Review of Economics and Statistic*s, MIT Press, vol. 79(4), pages 527-539, November.

[14] Bayudan, C. (2006). Wives' time allocation and intrahousehold power: evidence from the Philippines. *Applied Economics*, 38(7), 789-804.

[15] Becker, G. (1991). A Treatise on the Family. Harvard University Press, Cambridge, MA.

[16] Behrman, J. and Parker, S. (2011). The impact of the progresa/oportunidades conditional cash transfer program on health and related outcomes for the aging in mexico.

[17] Blau, D. and Goodstein, R. (2014). Commitment in the Household: Evidence from the Effect of Inheritances on the Labor Supply of Older Married Couples, mimeo, Department of Economics, Ohio State University.

[18] Blundell, R., P.-A. Chiappori and C. Meghir (2005), Collective labor supply with children, *Journal of Political Economy*, 113, 1277-1306.

[19] Blundell, R. and Robin, J. (1999). Estimation in Large and Disaggregated Demand Systems: An Estimator for Conditionally Linear Systems, *Journal of Applied Econometrics*, vol. 14(3), pages 209-32, May-June.

[20] Bourguignon, F., Browning, M. and P.A. Chiappori (2009). Efficient Intra-Household Allocations and Distribution Factors: Implication and Identification, *Review of Economic Studies*, 76, pp 503-28.

[21] Bourguignon, F., Browning, M. and P.A. Chiappori and Lechene, V. (1993). Intra-Household Allocation of Consumption: A Model and some Evidence from French Data. *Annales d'Economie et de Statistique*, No. 29, pp. 137-156.

[22] Bobonis, G. (2009). Is the Allocation of Resources within the Household Efficient? New Evidence from a Randomized Experiment. *Journal of Political Economy*, 117(3) (2009), 453-503.

[23] Bronson, M. (2015). Degrees are Forever: Marriage, Educational Investment, and Lifecycle Labor Decisions of Men and Women. Working Paper, UCLA.

[24] Browning M. and P.-A. Chiappori (1998). Efficient Intra-household Allocations: A General Characterization and Empirical Tests. *Econometrica*, vol. 66, pp. 1241-1278.

[25] Browning, M., Chiappori, P.-A., and Weiss, Y. (2014). Economics of the Family. Cambridge University Press.

[26] Browning M., Bourguignon, F., P.-A. Chiappori, and Lechene, V. (1994). Income and outcomes: a structural model of intrahousehold allocation. *Journal of Political Economy* 102:1067-96.

[27] Chiappori, P.-A. (1988), Rational household labor supply, *Econometrica*, 56, 63-89.

[28] Chiappori, P.-A. (1992), Collective Labor Supply and Welfare, *Journal of Political Economy*, vol. 100(3), pages 437-67, June.

[29] Chiappori, P.-A. and Ekeland, I. (2006). The Micro Economics of Group Behavior: General Characterization. *Journal of Economic Theory* 130:1-26.

[30] Chiappori, P.-A. and Ekeland, I. (2009). The Microeconomics of Efficient Group Behavior: Identification. *Econometrica* 77:763-799.

[31] Chiappori, P.-A., Fortin, B. Lacroix, G. (2002). Marriage Market, Divorce Legislation, and Household Labor Supply. *Journal of Political Economy* 110 (1): 37-72 (February).

[32] Chiappori, P.-A. and M. Mazzocco (2015). Static and intertemporal household decisions. mimeo.

[33] Cherchye, L., De Rock, B. and Vermeulen, F. (2009). Opening the Black Box of Intrahousehold Decision Making: Theory and Nonparametric Empirical Tests of General Collective Consumption Models. *Journal of Political Economy* 117:1074-1104.

[34] Cunha, F., Heckman, J. J., Lochner, L. J., and Masterov, D. V. (2006). Interpreting the evidence on life cycle skill formation. In E. A. Hanushek, and F. Welch (Eds.), *Handbook of the Economics of Education*, chap. 12. Amsterdam: North-Holland.

[35] Dauphin, A., Fortin, B. Lacroix, G. (2015). How Falsifiable is the Collective Model? A New Test with an Application to Monogamous and Bigamous Households in Burkina Faso, IZA Discussion Papers 9002, Institute for the Study of Labor (IZA).

[36] Dercon, S. and P. Krishnan (2000). In sickness and in health: Risk sharing within households in rural Ethiopia. *Journal of Political Economy* 108: 688-727.

[37] Donni, O. (2007). Collective female labour supply: theory and application. *The Economic Journal*, 117(516), 94-119.

[38] Donni, O. and Moreau, N. (2007). Collective Labor Supply A Single-Equation Model and Some Evidence from French Data. *Journal of Human Resources*, 42(1), 214-246.

[39] Duflo, E. (2003). Grandmothers and Granddaughters: Old Age Pension and Intra-Household Allocation in South Africa, *World Bank Economic Review*, 17, 1-25.

[40] Fortin, B. and Lacroix, G. (1997). A test of neoclassical and collective models of household labour supply, *Economic Journal*, 107, 933-955.

[41] Gertler, P. (2004). Do conditional cash transfers improve child health? evidence from progresa's control randomized experiment. *American economic review*, pages 336-341.

[42] Gemici, A., and Laufer, S. (2012). Marriage and Cohabitation. Working Paper, NYU.

[43] Goldstein, M. and Udry, C. (2008). The profits of power: Land rights and agricultural investment in Ghana. *Journal of Political Economy*, 116(6), 981-1022.

[44] LaFave, D. and Thomas, D. (2013). Extended Families and Child Well-being. Working Paper.

[45] Ligon, E. (2002). Dynamic Bargaining in Households (with an application to Bangladesh). Giannini Foundation Working Paper.

[46] Lise, J. and Yamada, K. (2014). Household Sharing and Commitment: Evidence from Panel Data on Individual Expenditures and Time Use. Institute for Fiscal Studies Working Papers.

[47] Lundberg, S., Pollak, R. (1993). Separate Spheres Bargaining and the Marriage Market. *Journal of Political Economy* 101 (6): 988-1010.

[48] Lundberg, S., Pollak, R., and Wales, T. (1997). Do Husbands and Wives Pool Their Resources? Evidence from the UK Child Benefit, *Journal of Human Resources* 32(3), 463-480.

[49] Matsuda, T. (2006). Linear approximations to the quadratic almost ideal demand system, *Empirical Economics*, Volume 31, Issue 3, pp 663-675.

[50] Mazzocco, M. (2007). Household Intertemporal Behaviour: A Collective Characterization and a Test of Commitment. *Review of Economic Studies* 74:857-895.

[51] Mazzocco, M., Ruiz, C. and Yamaguchi, S. (2014). Labor Supply and Household Dynamics. *American Economic Review: Papers & Proceedings* 104 (5): 354-59.

[52] Moschini G. (1995) Units of measurement and the Stone index in demand system estimation. *American Journal of Agricultural Economics*, 77:63-68.

[53] Naidoo, J. (2015). The power of tests for pareto efficiency within the family. Mimeo, University of Chicago.

[54] Quisumbing, A. and Maluccio, J. (2003). Resources at Marriage and Intrahousehold Allocation: Evidence from Bangladesh, Ethiopia, Indonesia, and South Africa, *Oxford Bulletin of Economics and Statistics*, vol. 65(3), pages 283-327.

[55] Robinson, J. (2012). Limited insurance within the household: Evidence from a field experiment in Kenya. *American Economic Journal: Applied Economics* 4: 140-164.

[56] Rubalcava, L., Teruel, G. and Thomas, D. (2008). Investments, Time Preferences, and Public Transfers Paid to Women. *Economic Development and Cultural Change*, 57(3): 507-38.

[57] Schultz, T. P. (2004). School subsidies for the poor: evaluating the mexican progresa poverty program. *Journal of Development Economics*, 74(1):199-250.

[58] Thomas, D. D. Contreras, and E. Frankenberg (2002). Distribution of power within the household and child health. *RAND*, Santa Monica, California.

[59] Tommasi, D. (2015). How Cash Transfers Improve Child Development, Working Papers ECARES 2015-19, ULB - Universite Libre de Bruxelles.

[60] Tommasi, D. and Wolf, A. (2016). An empirical bayes method to estimate resource shares of household members. mimeo.

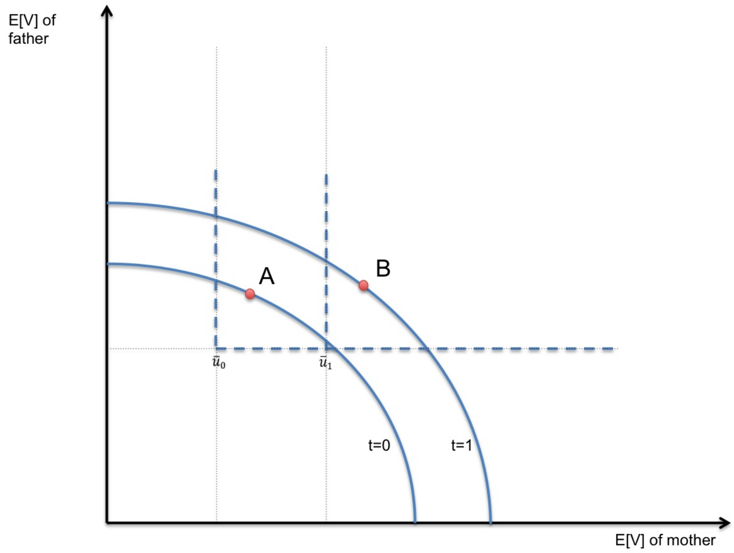
[61] Udry, C. (1996). Gender, agricultural production, and the theory of the household. *Journal of Political Economy* 104: 1010-1046.

[62] Vermeulen, F. (2005). And the winner is... An empirical evaluation of unitary and collective labour supply models. *Empirical Economics*, 30(3), 711=734.

[63] Voena, A. (2015). Yours, Mine and Ours: Do Divorce Laws Affect the Intertemporal Behavior of Married Couples? *American Economic Review, Forthcoming.*

Appendix

Figure 1: Graphical illustration of the LIC model



Notes: This figure provides a graphical illustration of how the Limited-commitment Intertemporal Collective (LIC) model works.

Table 1: Literature review of tests on efficiency

Paper	**Country**	**Subject**	**Type of test**	**Conclusion**
Adamowicza et al. (2013)	US	consumer goods	other test	fails to reject
Angelucci and Garlick (2015)	Mexico	consumer goods	z-conditional	fails to reject only for old couples
Attanasio and Lechene (2014)	Mexico	consumer goods	z-conditional	fails to reject
Bayudan (2006)	Philippines	time use	proportionality	fails to reject
Blau and Goodstein (2014)	US	labor supply	commitment	rejects
Bobonis (2009)	Mexico	consumer goods	z-conditional	fails to reject
Bourguignon et al. (1993)	France	consumer goods	proportionality	fails to reject
Browning et al. (1994)	Canada	consumer goods	proportionality	fails to reject
Browning and Chiappori (1998)	Canada	consumer goods	proportionality	fails to reject
Chiappori et al. (2002)	US	labor supply	proportionality	fails to reject
Donni (2007)	France	consumer goods	proportionality	fails to reject
Donni and Moreau (2007)	France	labor supply	proportionality	fails to reject
Fortin and Lacroix (1997)	Canada	labor supply	proportionality	fails to reject
Goldstein and Udry (2008)	Ghana	land use	other test	rejects
LaFave and Thomas (2013)	Indonesia	child well being	proportionality	fails to reject
Lise and Yamada (2014)	Japan	risk-sharing, credit	commitment	rejects
Mazzocco (2007)	US	risk-sharing, credit	commitment	rejects
Thomas et al. (2002)	Indonesia	child health	proportionality	fails to reject
Vermeulen (2005)	Netherlands	labor supply	proportionality	fails to reject
Voena (2010)	US	risk-sharing, credit	commitment	rejects
Udry (1996)	Burkina Faso	land use	other test	rejects

Notes: Selected empirical studies of allocation within the family.

Table 2: Test of commitment, full sample

Specification	(1)	(2)	(3)	(4)	(5)	(6)
Joint test of:				*1998*		
Treatment	0.08	0.05	0.03	0.00	0.06	0.07
Network		0.00	0.00	0.04	0.00	0.22
(1 + Edu_m) / (1 + Edu_f)			0.02	0.65		
Treatment*(1 + Edu_m) / (1 + Edu_f)				0.28		
Treatment*Network				0.00		0.98
$Edu_m = Edu_f$					0.38	0.09
$Edu_m > Edu_f$					0.07	0.07
Treatment*($Edu_m = Edu_f$)						0.32
Treatment*($Edu_m > Edu_f$)						0.37
Observations	6,153	5,125	5,125	5,125	5,125	5,125
Joint test of:				*1999*		
Treatment	0.00	0.00	0.00	0.00	0.00	0.00
Network		0.01	0.01	0.00	0.03	0.00
(1 + Edu_m) / (1 + Edu_f)			0.00	0.05		
Treatment*(1 + Edu_m) / (1 + Edu_f)				0.00		
Treatment*Network				0.00		0.01
$Edu_m = Edu_f$					0.03	0.05
$Edu_m > Edu_f$					0.12	0.09
Treatment*($Edu_m = Edu_f$)						0.04
Treatment*($Edu_m > Edu_f$)						0.30
Observations	5,739	4,932	4,932	4,932	4,932	4,932
Joint test of:				*Pooled*		
Treatment	0.00	0.00	0.00	0.00	0.00	0.04
Network		0.00	0.00	0.00	0.00	0.00
(1 + Edu_m) / (1 + Edu_f)			0.00	0.00		
Treatment*(1 + Edu_m) / (1 + Edu_f)				0.00		
Treatment*Network				0.00		0.21
$Edu_m = Edu_f$					0.32	0.17
$Edu_m > Edu_f$					0.13	0.03
Treatment*($Edu_m = Edu_f$)						0.08
Treatment*($Edu_m > Edu_f$)						0.13
Observations	11,892	10,057	10,057	10,057	10,057	10,057

Notes: The table reports the p-values of the joint significance tests of treatment for 6 different specifications of system (16), separately for 1998, 1999, and the pooled sample. The empirical test that we run corresponds to the restriction implied by equation (13). The blue boxes indicate when the test yields p-values below the 10% level, hence when we reject the null hypothesis that the joint test is not significant. The preferred third distribution factor is the the education ratio, computed as (1 + Edu_m) / (1 + Edu_f). The full set of parameter estimates of the preferred specification (3) can be found in the Appendix. The parameter estimates of all the other specifications are available upon request.

Table 3: Test of commitment, sub-samples

Specification	(7) $\text{Edu}_m < \text{Edu}_f$	(8) $\text{Edu}_m = \text{Edu}_f$	(9) $\text{Edu}_m > \text{Edu}_f$	(10) $18 < \text{Age}_f < 30$	(11) $30 \leq \text{Age}_f < 40$	(12) $40 \leq \text{Age}_f < 50$	(13) $\text{Age}_f \geq 50$
Joint test of:				*1998*			
Treatment	0.11	0.44	0.88	0.29	0.08	0.89	0.92
Network	0.20	0.05	0.05	0.06	0.04	0.73	0.55
$\text{Edu}_m = \text{Edu}_f$				0.03	0.99	0.22	0.48
$\text{Edu}_m > \text{Edu}_f$				0.00	0.73	0.37	0.53
Observations	1,341	2,806	978	1,498	2,098	954	575
Joint test of:				*1999*			
Treatment	0.00	0.00	0.00	0.00	0.00	0.01	0.23
Network	0.48	0.40	0.04	0.32	0.01	0.30	0.08
$\text{Edu}_m = \text{Edu}_f$				0.19	0.65	0.37	0.91
$\text{Edu}_m > \text{Edu}_f$				0.55	0.80	0.82	0.60
Observations	1,289	2,718	925	1,367	2,047	926	592
Joint test of:				*Pooled*			
Treatment	0.23	0.00	0.01	0.01	0.00	0.00	0.60
Network	0.13	0.00	0.01	0.39	0.00	0.47	0.02
$\text{Edu}_m = \text{Edu}_f$				0.00	0.83	0.58	0.42
$\text{Edu}_m > \text{Edu}_f$				0.15	0.82	0.53	0.22
Observations	2,630	5,524	1,903	2,865	4,145	1,880	1,167

Notes: The table reports the p-values of the joint significance tests of treatment for 7 different specifications of system (16), separately for 1998, 1999, and the pooled sample. The empirical test that we run corresponds to the restriction implied by equation (13). The blue boxes indicate when the test yields p-values below the 10% level, hence when we reject the null hypothesis that the joint test is not significant. The full set of parameter estimates of all specifications are available upon request. We choose to use this specification choice of education to make the first three tests comparable with the last four. We obtain the same qualitative results if we run specifications (10)-(13) with the education ratio $(1 + \text{Edu}_m) \;/\; (1 + \text{Edu}_f)$.

Table 4: Test of efficiency: "all or nothing" condition, full sample

Budget shares	w1	w2	w3	w4	w5
Distribution factors:			*1998*		
Treatment	0.020**	0.004	-0.012**	-0.016**	-0.004
	(0.009)	(0.006)	(0.005)	(0.008)	(0.003)
Network	-0.013*	-0.005	0.011***	0.013**	0.006*
	(0.007)	(0.004)	(0.004)	(0.005)	(0.004)
Education	0.085***	-0.004	-0.029**	-0.014	0.038***
	(0.030)	(0.015)	(0.014)	(0.026)	(0.014)
Observations	5,125	5,125	5,125	5,125	5,125
Distribution factors:			*1999*		
Treatment	-0.049***	-0.021**	0.021***	0.007	-0.041***
	(0.008)	(0.009)	(0.005)	(0.006)	(0.003)
Network	0.013**	0.003	-0.000	-0.002	0.013***
	(0.006)	(0.003)	(0.003)	(0.005)	(0.004)
Education	0.154***	0.023	-0.018	-0.112***	0.047***
	(0.030)	(0.015)	(0.014)	(0.027)	(0.015)
Observations	4,932	4,932	4,932	4,932	4,932
Distribution factors:			*Pooled*		
Treatment	-0.015*	-0.011**	0.001	-0.003	-0.027***
	(0.008)	(0.005)	(0.004)	(0.006)	(0.002)
Network	0.003	0.003	0.006**	0.002	0.013***
	(0.005)	(0.003)	(0.003)	(0.004)	(0.002)
Education	0.094***	0.001	-0.004	-0.051***	0.040***
	(0.019)	(0.008)	(0.011)	(0.017)	(0.010)
Observations	10,057	10,057	10,057	10,057	10,057

Notes: We report only the parameter estimates and the standard deviation attached to the distribution factors. The variable Education refers to the education ratio computed as (1 + Edu_m) / (1 + Edu_f). The complete set of parameter estimates can be found in the Appendix. *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Test of efficiency: z-conditional demand test, network as conditioning distribution factor

Dependent variable	w1	w1	w3	w3	w4	w4
Distribution factors:			*1998*			
Treatment	0.011	0.006	0.010	-0.002	-0.003	0.007
	(0.013)	(0.010)	(0.006)	(0.006)	(0.010)	(0.010)
Education	0.011	0.028	0.057**	-0.014	0.051	0.051**
	(0.030)	(0.019)	(0.027)	(0.013)	(0.035)	(0.023)
Conditioning good	w3	w4	w1	w4	w1	w3
Observations	5,125	5,125	5,125	5,125	5,125	5,125
Joint test of:						
Treatment	0.41	0.51	0.11	0.78	0.76	0.46
Education	0.71	0.14	0.03	0.31	0.14	0.03
Dependent variable	w1	w1	w4	w4	w5	w5
Distribution factors:			*1999*			
Treatment	-0.042**	-0.042*	-0.021	0.009	-0.026	0.009
	(0.016)	(0.021)	(0.019)	(0.022)	(0.023)	(0.016)
Education	0.117	0.050	0.048	-0.015	0.076	0.234
	(0.142)	(0.035)	(0.064)	(0.036)	(0.076)	(0.238)
Conditioning good	w4	w5	w1	w5	w1	w4
Observations	4,932	4,932	4,932	4,932	4,932	4,932
Joint test of:						
Treatment	0.01	0.05	0.25	0.67	0.27	0.57
Education	0.41	0.15	0.45	0.69	0.32	0.33
Dependent variable	w1	w1	w4	w4	w5	w5
Distribution factors:			*Pooled*			
Treatment	-0.000	-0.008	0.017	-0.013	-0.053***	0.024***
	(0.010)	(0.012)	(0.018)	(0.013)	(0.020)	(0.007)
Education difference	0.031	0.050**	-0.172	0.007	0.525***	-0.059
	(0.095)	(0.023)	(0.117)	(0.024)	(0.140)	(0.056)
Conditioning good	w4	w5	w1	w5	w1	w4
Observations	10,057	10,057	10,057	10,057	10,057	10,057
Joint test of:						
Treatment	1.00	0.49	0.33	0.29	0.01	0.00
Education	0.74	0.03	0.14	0.78	0.00	0.29

Notes: We report only the parameter estimates and the standard deviation attached to the two remaining distribution factors on which we test the collective model. The variable Education refers to the education ratio computed as (1 + Edu_m) / (1 + Edu_f). *** p<0.01, ** p<0.05, * p<0.1.

Table 6: Test of efficiency: "all or nothing" condition, control group

Budget shares	w1	w2	w3	w4	w5
Distribution factors:			*1998*		
Network	-0.006	0.001	0.005	0.005	0.006
	(0.009)	(0.005)	(0.006)	(0.008)	(0.006)
Education	0.071	-0.005	-0.019	-0.016	0.031
	(0.048)	(0.023)	(0.022)	(0.028)	(0.024)
Observations	1,949	1,949	1,949	1,949	1,949
Distribution factors:			*1999*		
Network	0.019**	0.007*	-0.000	-0.009	0.017***
	(0.010)	(0.004)	(0.005)	(0.008)	(0.006)
Education	0.126**	0.002	-0.031*	-0.043	0.054**
	(0.049)	(0.020)	(0.018)	(0.037)	(0.024)
Observations	1,858	1,858	1,858	1,858	1,858
Distribution factors:			*Pooled*		
Network	0.007	0.005	0.002	-0.003	0.012***
	(0.008)	(0.004)	(0.004)	(0.006)	(0.004)
Education	0.071**	-0.008	-0.001	-0.018	0.044**
	(0.033)	(0.014)	(0.014)	(0.023)	(0.019)
Observations	3,807	3,807	3,807	3,807	3,807

Notes: We report only the parameter estimates and the standard deviation attached to the distribution factors. The variable Education refers to the education ratio computed as (1 + Edu_m) / (1 + Edu_f). *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 7: Test of efficiency: z-conditional demand test, control group

Dependent variable	w1	w5
Distribution factors:	*1999*	
Education	-0.033	0.053
	(0.060)	(0.062)
Conditioning good	w5	w1
Observations	1,858	1,858
Joint test of:		
Education	0.58	0.39
Dependent variable	w1	w5
Distribution factors:	*Pooled*	
Education	0.055	0.007
	(0.040)	(0.049)
Conditioning good	w5	w1
Observations	3,807	3,807
Joint test of:		
Education	0.17	0.89

Notes: We report only the parameter estimates and the standard deviation attached to the remaining distribution factor on which we test the collective model. The variable Education refers to the education ratio computed as (1 + Edu_m) / (1 + Edu_f). *** p<0.01, ** p<0.05, * p<0.1.

Table 8: Test of efficiency: "all or nothing" condition, treatment group

Budget shares	w1	w2	w3	w4	w5
Distribution factors:			*1998*		
Network	-0.019**	-0.009	0.015***	0.023***	0.010*
	(0.009)	(0.007)	(0.006)	(0.008)	(0.006)
Education	0.073*	-0.011	-0.026	0.001	0.037
	(0.038)	(0.019)	(0.020)	(0.042)	(0.024)
Observations	3,176	3,176	3,176	3,176	3,176
Distribution factors:			*1999*		
Network	0.006	-0.000	-0.001	0.002	0.007
	(0.008)	(0.003)	(0.004)	(0.007)	(0.005)
Education	0.178***	0.015	-0.002	-0.138***	0.053**
	(0.035)	(0.019)	(0.021)	(0.034)	(0.021)
Observations	3,074	3,074	3,074	3,074	3,074
Distribution factors:			*Pooled*		
Network	0.003	0.003	0.005	0.004	0.015***
	(0.006)	(0.003)	(0.003)	(0.006)	(0.002)
Education	0.096***	0.000	0.008	-0.067***	0.037**
	(0.029)	(0.012)	(0.016)	(0.026)	(0.017)
Observations	6,250	6,250	6,250	6,250	6,250

Notes: We report only the parameter estimates and the standard deviation attached to the distribution factors. The variable Education refers to the education ratio computed as $(1 + \text{Edu}_m) / (1 + \text{Edu}_f)$. *** p<0.01, ** p<0.05, * p<0.1.

Table 9: Test of efficiency: z-conditional demand test, treatment group

Dependent variable	w1	w1	w3	w3	w4	w4
Distribution factors:			*1998*			
Education	-0.010	-0.007	0.022	-0.011	0.042	0.066**
	(0.032)	(0.027)	(0.023)	(0.020)	(0.037)	(0.026)
Conditioning good	w3	w4	w1	w4	w1	w3
Observations	3,176	3,176	3,176	3,176	3,176	3,176
Joint test of:						
Education	0.76	0.80	0.34	0.57	0.26	0.01
Dependent variable	w1	w1	w4	w4	w5	w5
Distribution factors:			*1999*			
Education	0.081	0.036	-0.011	-0.053	0.062	0.038
	(0.223)	(0.079)	(0.113)	(0.058)	(0.146)	(0.503)
Conditioning good	w4	w5	w1	w5	w1	w4
Observations	3,074	3,074	3,074	3,074	3,074	3,074
Joint test of:						
Education	0.72	0.65	0.92	0.36	0.67	0.94
Dependent variable	w1	w1	w4	w4	w5	w5
Distribution factors:			*Pooled*			
Education	-0.093	0.020	0.013	-0.019	0.071	-0.100
	(0.067)	(0.035)	(0.070)	(0.024)	(0.060)	(0.083)
Conditioning good	w4	w5	w1	w5	w1	w4
Observations	6,250	6,250	6,250	6,250	6,250	6,250
Joint test of:						
Education	0.17	0.57	0.86	0.42	0.24	0.23

Notes: We report only the parameter estimates and the standard deviation attached to the remaining distribution factor on which we test the collective model. The variable Education refers to the education ratio computed as (1 + Edu_m) / (1 + Edu_f). *** p<0.01, ** p<0.05, * p<0.1.

A: Online Appendix

A1 Statistical discussion of the results of the test of commitment

The results of the test of commitment is Section 5.1 indicate that commitment is broken but it requires a large exogenous shock to be picked up by the test procedure. The effect of the treatment is not significant in smaller samples in 1998 because, at the beginning of the program, the Minimum Detectable Effect (MDE) size is small (Bloom, 2006). This is logically obvious as the only difference between the full and the sub-samples is the sample size n. The formula of the MDE is:

$$MDE = M_{j-G-2}\sqrt{\frac{\rho(1-R_2^2)}{P(1-P)J} + \frac{(1-\rho)(1-R_1^2)}{P(1-P)Jn}} \tag{A1}$$

where:

- Multiplier for two-tailed test: with $J - g * -2$ degrees of freedom;
- J: the total number of clusters;
- $g*$: the number of group covariates used;
- $\rho = \frac{\tau^2}{\tau^2+\sigma^2}$ is the unconditional intra-class coefficient (ICC);
- τ^2: Level-2 (between group-level) variance in the unconditional model (without any covariates);
- σ^2: Level-1 (individual-level) variance in the unconditional model; is the proportion of variance in the outcome measure occurring at level one that is explained by covariates, X;
- $R_1^2 = 1 - (\frac{\sigma^2_{|X}}{\sigma^2})$: is the proportion of the within group variance (at level two) that is explained by the covariates, W;
- P: the proportion of this sample assigned to the treatment group ($\frac{J_T}{J}$);

As the sample size n decreases, the MDE increases, which means that we can only identify the effect of the treatment if this effect is large. In our specific case, this simply means that the eligibility to the cash transfers *does* bring the household to renegotiate the initial contract of resource allocation, but the shift is very small.

A2 Theoretical discussion of the results of the test of commitment

One way to interpret the results of the test of commitment explained in the previous Section may be the following. At the beginning of the program, the shift in the budget structure exists but it is small because a household may take time to change behavior: e.g. reluctance, social norms, lagged habits, lagged shift in bargaining power, learning. In principle these ideas can be incorporated in our Limited-commitment Intertemporal Collective (LIC) model of Section 2. Essentially the task is to specify a functional (*distributed-lag*) form for the inter temporal bargaining power within Mazzocco (2007)'s framework, which is left unspecified.

Consider the LIC model that we outlined in Section 2, this time with a 3-period horizon. Like before, in the first period the household is formed, and in the second period a new policy is installed. Now, in the third period, the policy is established and eligible households keep receiving the cash transfers associated with it. Although the design of the policy does not change with respect to period two, we may introduce a third period to distinguish the short run, the *installation* of the policy, from the medium/long run, the *establishment* of the policy. The resources of the family are derived from a household-level endowment A_1, realized at the time of marriage, total household earnings x_t, realized at each point in time, and an endowment entitled to member m which is realized from period 2 onward. The endowment is realized at the beginning of period 2 as well as at the beginning of period 3, before period-2 and -3 choices are made.

The advantage of this interpretation is to capture the idea that a distribution factor may have *time-differentiated effects* on the pattern of the outside option for member i. For instance, take our example of the cash transfers entitled to the mother. The grant becomes available to the mother at period $t = 2$, after the household is formed, and continues to be paid off in period $t = 3$. For a variety of reasons, such as habits, social norms, learning, etc, even if the grant becomes available to the mother in period two, in the short run (or *installation* of the policy) her outside option may be still not perceived as different from the outside option at period $t = 1$. It may be that it is only when the policy is *established*, that the (perceived) outside option becomes significantly different from the one at $t = 1$. We can capture all these ideas in a simple way as follows:

$$\bar{u}_{it}(Z) = \beta_0 + \beta_1(z^*_{it} = 1) + \beta_2(z^*_{it-1} = 1) + v \tag{A2}$$

where $\beta_0 = \bar{u}_{i1}(Z)$ is the value of the outside option in period 1, at the time of the household formation, β_1 and β_2 capture the marginal increase in the value of the outside option as the cash transfers becomes available in, respectively, period t and $t-1$, and v is a random error.

Note that, since the outside option is unobserved in practice, the parameters $(\beta_0, \beta_1, \beta_2)$ cannot be actually estimated, hence this interpretation cannot be tested formally as we lack the necessary data, but it is an interesting venue that should be further explored. Hence the advantage of making such a restriction is only interpretative for the results of the empirical section. This interpretation of the pattern of the outside option becomes interesting when, for instance, even if the design of the policy does not change over time, its effects on the intra-household decision process become significant only later in time, which is indeed the case for the PROGRESA grant.

A3 QAIDS versus $\ell-$QAIDS

Concerning the implementation of the z-conditional demand test, the main difference between the present paper and Bobonis (2009), Attanasio and Lechene (2014) and Angelucci and Garlick (2015), is the different specification of the demand system. QAIDS and $\ell-$QAIDS differ in the construction of the price indices. In the former:

$$\begin{aligned} \ln a(\mathbf{p}) &= \alpha_0 + \sum_{i=1}^{n} \alpha_i \ln p_i + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \gamma_{ij} \ln p_i \ln p_k \\ \ln b(\mathbf{p}) &= \beta_0 + \sum_{i=1}^{n} \beta_i \ln p_i \end{aligned} \tag{A3}$$

Whereas in the latter, a version of the Stone index (called Laspeyres index) is often used in place of (A3):

$$\ln P^* = \sum_{i=1}^{n} \bar{w}_i \ln(p_i) \tag{A4}$$

where $\bar{w}$ is the mean (or median) budget share among the n commodities. The $\ell-$QAIDS model with the Laspeyres index can be denoted as follows:

$$w_k = \hat{\alpha}_k + \sum_{j=1}^{k} \gamma_{kj} \ln p_j + \beta_k \ln \hat{x} + \lambda_k \ln \hat{x}^2 + \hat{\epsilon}_k \tag{A5}$$

where the implicit assumptions made are:

- $\hat{x} = \frac{x}{\hat{a}(\mathbf{p})}$
- $\hat{\alpha}_k = \alpha_k - \beta_k \alpha_k - \frac{\lambda_k}{b(\mathbf{p})} \alpha_k$
- $\hat{\epsilon}_k = \epsilon_k - \beta_k \big[\ln(\phi) - E[\ln(\phi)]\big] - \frac{\lambda_k}{b(\mathbf{p})} \big[\ln(\phi^2) - E[\ln(\phi^2)]\big]$

These are quite strong assumptions and, as it is widely known, applying (A4) instead of (A3) can seriously bias elasticity estimates partly because this price index is influenced by changes in units of measurement (Alston et al., 1994; Moschini, 1995; Asche and Wessells, 1997; Matsuda, 2006).

A4 Additional results

Table A1: Parameter estimates

	1998				1999				Pooled			
	w1	w2	w3	w4	w1	w2	w3	w4	w1	w2	w3	w4
Intercept:	0.108*	0.154***	0.223***	0.391***	0.291***	0.380***	0.015	0.373***	0.250***	0.247***	0.138***	0.369***
	(0.056)	(0.037)	(0.028)	(0.047)	(0.042)	(0.077)	(0.038)	(0.037)	(0.032)	(0.032)	(0.022)	(0.032)
Expenditure:	-0.051	-0.100	0.050	0.204***	0.261***	0.118**	-0.080**	-0.023	0.154***	0.053	-0.003	0.060
	(0.095)	(0.065)	(0.046)	(0.078)	(0.049)	(0.057)	(0.038)	(0.043)	(0.043)	(0.038)	(0.029)	(0.037)
Prices:	-0.095***	0.041**	0.040**	0.025	-0.124***	0.058***	0.016	0.089***	-0.119***	0.009	0.028***	0.101***
	(0.034)	(0.021)	(0.017)	(0.030)	(0.018)	(0.019)	(0.011)	(0.013)	(0.012)	(0.008)	(0.006)	(0.010)
		-0.008	0.033*	-0.087***		-0.276***	0.152***	-0.021		-0.096***	0.081***	-0.028***
		(0.025)	(0.018)	(0.024)		(0.083)	(0.034)	(0.018)		(0.024)	(0.011)	(0.011)
			-0.058***	0.029*			-0.078***	-0.049***			-0.008	-0.046***
			(0.018)	(0.016)			(0.019)	(0.010)			(0.009)	(0.006)
				0.124***				0.051***				0.043***
				(0.035)				(0.016)				(0.012)
Expenditure squared:	0.066	0.010	-0.044	-0.043	0.006	-0.003	0.027	-0.045**	0.009	-0.003	0.011	-0.024
	(0.054)	(0.037)	(0.029)	(0.048)	(0.026)	(0.013)	(0.022)	(0.023)	(0.022)	(0.013)	(0.017)	(0.018)
Number of children in primary:	0.023***	0.010***	-0.011***	-0.022***	-0.013***	-0.007	0.006**	-0.005	0.005	-0.002	-0.006***	-0.012***
	(0.005)	(0.004)	(0.003)	(0.004)	(0.004)	(0.006)	(0.003)	(0.003)	(0.004)	(0.003)	(0.002)	(0.003)
Number of children in secondary:	-0.136**	-0.052	0.087***	0.056	0.111*	-0.017	0.008	0.053	-0.097**	-0.030	0.112***	0.077**
	(0.063)	(0.032)	(0.026)	(0.046)	(0.060)	(0.041)	(0.025)	(0.044)	(0.049)	(0.030)	(0.019)	(0.033)
Number of younger children:	-0.004	-0.004	0.003	0.000	0.011*	-0.004	0.000	0.002	-0.005	-0.004	0.008***	0.004
	(0.006)	(0.003)	(0.003)	(0.005)	(0.006)	(0.003)	(0.002)	(0.004)	(0.005)	(0.003)	(0.002)	(0.003)
Education of the mother:	-0.030***	-0.003	0.016***	0.009	-0.047***	-0.009*	0.006	0.039***	-0.028***	-0.002	0.003	0.018***
	(0.010)	(0.005)	(0.005)	(0.009)	(0.009)	(0.005)	(0.005)	(0.009)	(0.006)	(0.003)	(0.004)	(0.005)
Education of the father:	0.017	-0.006	-0.001	0.001	0.041***	0.002	-0.004	-0.035***	0.024***	-0.003	0.002	-0.015**
	(0.010)	(0.005)	(0.005)	(0.009)	(0.010)	(0.005)	(0.005)	(0.009)	(0.007)	(0.003)	(0.004)	(0.006)
Father is indigenous:	-0.029***	-0.027***	0.021***	0.023***	0.019***	-0.003	-0.005	-0.007	0.004	-0.007*	0.008**	0.002
	(0.009)	(0.006)	(0.005)	(0.007)	(0.007)	(0.004)	(0.003)	(0.006)	(0.006)	(0.004)	(0.004)	(0.006)
Age of the father:	0.002***	0.001*	-0.001***	-0.001**	-0.002**	-0.000	0.000	-0.001**	0.001	0.000	-0.001***	-0.001***
	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
Size of the town:	0.000	-0.000	0.000	-0.000	0.000	0.000*	-0.000	-0.000*	0.000**	0.000	0.000	-0.000*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Relatives eat in the house:	-0.008**	-0.000	0.002	0.002	-0.002	-0.001	0.000	0.000	-0.006**	-0.001	0.002	0.001
	(0.004)	(0.002)	(0.004)	(0.004)	(0.003)	(0.001)	(0.001)	(0.002)	(0.003)	(0.001)	(0.001)	(0.002)
Member of the family eats out:	0.031*	-0.001	-0.017**	-0.022	-0.004	-0.003	0.004	0.005	0.006	-0.008**	-0.003	-0.001
	(0.018)	(0.008)	(0.008)	(0.016)	(0.009)	(0.004)	(0.004)	(0.007)	(0.008)	(0.003)	(0.004)	(0.007)
Treatment:	0.020**	0.004	-0.012**	-0.016**	-0.049***	-0.021**	0.021***	0.007	-0.015*	-0.011**	0.001	-0.003
	(0.009)	(0.006)	(0.005)	(0.008)	(0.008)	(0.009)	(0.005)	(0.006)	(0.008)	(0.005)	(0.004)	(0.006)
Network:	-0.013*	-0.005	0.011***	0.013**	0.013**	0.003	-0.000	-0.002	0.003	0.003	0.006**	0.002
	(0.007)	(0.004)	(0.004)	(0.005)	(0.006)	(0.003)	(0.003)	(0.005)	(0.005)	(0.003)	(0.003)	(0.004)
Education difference:	0.085***	-0.004	-0.029**	-0.014	0.154***	0.023	-0.018	-0.112***	0.094***	0.001	-0.004	-0.051***
	(0.030)	(0.015)	(0.014)	(0.026)	(0.030)	(0.015)	(0.014)	(0.027)	(0.019)	(0.008)	(0.011)	(0.017)
Residuals:	0.520	0.050	-0.325	-0.421	-0.090	-0.394**	0.573**	-0.430	-0.149	-0.312**	0.336*	-0.175
	(0.527)	(0.318)	(0.273)	(0.455)	(0.324)	(0.169)	(0.275)	(0.311)	(0.273)	(0.150)	(0.189)	(0.219)
	0.121	0.006	-0.059	0.099*	0.133*	0.012	-0.096**	0.102**	0.151***	-0.011	-0.072**	0.086**
	(0.087)	(0.036)	(0.049)	(0.058)	(0.070)	(0.033)	(0.047)	(0.049)	(0.057)	(0.026)	(0.036)	(0.039)
	-0.040	0.003	0.023	0.028	-0.015	0.024	-0.048*	0.050	0.003	0.022	-0.034*	0.019
	(0.057)	(0.034)	(0.029)	(0.048)	(0.034)	(0.017)	(0.028)	(0.032)	(0.028)	(0.016)	(0.019)	(0.023)
	-0.002*	-0.000	0.001	-0.001*	-0.002***	-0.000	0.001***	-0.001*	-0.002***	-0.000	0.001***	-0.001**
	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
	0.141**	0.054*	-0.087***	-0.062	-0.098	0.015	-0.007	-0.061	0.107**	0.030	-0.112***	-0.084**
	(0.063)	(0.032)	(0.026)	(0.046)	(0.060)	(0.041)	(0.026)	(0.044)	(0.048)	(0.030)	(0.019)	(0.033)
	0.005	-0.000	0.001	-0.003	-0.010**	0.003	-0.000	0.006	-0.002	0.001	0.001	0.001
	(0.004)	(0.002)	(0.003)	(0.003)	(0.005)	(0.002)	(0.003)	(0.004)	(0.003)	(0.002)	(0.002)	(0.003)
Observations	5,125	5,125	5,125	5,125	4,932	4,932	4,932	4,932	10,057	10,057	10,057	10,057
Joint test of:												
Treatment	0.03				0.00				0.00			
Network	0.00				0.01				0.00			
Education	0.02				0.00				0.00			

Notes: *** p<0.01, ** p<0.05, * p<0.1.

Table A2: Test of efficiency: z-conditional demand test, education as conditioning good

Dependent variable	w1	w1	w3	w3	w4	w4
Distribution factors:			*1998*			
Treatment	0.013	0.039	-0.002	0.001	-0.016*	-0.016
	(0.018)	(0.031)	(0.007)	(0.013)	(0.009)	(0.013)
Network	-0.009	-0.032	0.006	0.003	0.011*	0.011
	(0.016)	(0.025)	(0.005)	(0.012)	(0.006)	(0.011)
Conditioning good	w3	w4	w1	w4	w1	w3
Observations	5,125	5,125	5,125	5,125	5,125	5,125
Joint test of:						
Treatment	0.49	0.21	0.73	0.96	0.07	0.23
Network	0.55	0.21	0.23	0.82	0.09	0.32
Distribution factors:			*1999*			
Treatment	-0.013	-0.036*	0.014	0.029**	-0.006	0.014
	(0.022)	(0.020)	(0.012)	(0.012)	(0.014)	(0.027)
Network	0.016**	0.009	0.004	0.003	0.002	-0.006
	(0.007)	(0.006)	(0.006)	(0.005)	(0.005)	(0.007)
Conditioning good	w3	w4	w1	w4	w1	w3
Observations	4,932	4,932	4,932	4,932	4,932	4,932
Joint test of:						
Treatment	0.55	0.07	0.22	0.02	0.66	0.61
Education	0.03	0.15	0.50	0.57	0.76	0.35
Distribution factors:			*Pooled*			
Treatment	0.001	0.000	-0.010*	-0.011*	-0.008	-0.025**
	(0.009)	(0.010)	(0.005)	(0.006)	(0.009)	(0.012)
Network	-0.017	0.001	0.008***	0.009***	0.004	0.002
	(0.020)	(0.004)	(0.003)	(0.003)	(0.004)	(0.020)
Conditioning good	w3	w4	w1	w4	w1	w3
Observations	10,057	10,057	10,057	10,057	10,057	10,057
Joint test of:						
Treatment	0.90	0.99	0.06	0.05	0.36	0.04
Education	0.39	0.76	0.00	0.00	0.29	0.92

Notes: We report only the parameter estimates and the standard deviation attached to the two remaining distribution factors on which we test the collective model. The variable Education refers to the education ratio computed as (1 + Edu_m) / (1 + Edu_f). *** p<0.01, ** p<0.05, * p<0.1.

Table A3: Test of efficiency: "all or nothing" condition, older heads

Budget shares	w1	w2	w3	w4	w5
Distribution factors:			*1998*		
Treatment	0.014	-0.006	-0.010	-0.006	-0.009
	(0.012)	(0.007)	(0.007)	(0.010)	(0.007)
Network	-0.011	-0.001	0.009*	0.009	0.007
	(0.011)	(0.005)	(0.005)	(0.008)	(0.006)
Education	0.030	-0.039	-0.047	0.044	-0.013
	(0.049)	(0.026)	(0.029)	(0.044)	(0.029)
Observations	2,436	2,436	2,436	2,436	2,436
Distribution factors:			*1999*		
Treatment	-0.054***	-0.005	0.015**	0.007	-0.037***
	(0.010)	(0.011)	(0.006)	(0.007)	(0.006)
Network	0.011	-0.003	0.004	-0.004	0.007
	(0.009)	(0.004)	(0.005)	(0.007)	(0.005)
Education	0.134***	0.016	0.001	-0.074*	0.077***
	(0.045)	(0.024)	(0.023)	(0.038)	(0.026)
Observations	2,448	2,448	2,448	2,448	2,448
Distribution factors:			*Pooled*		
Treatment	-0.022**	-0.016***	0.003	0.002	-0.034***
	(0.009)	(0.006)	(0.005)	(0.006)	(0.005)
Network	0.005	-0.000	0.006**	-0.000	0.011***
	(0.006)	(0.003)	(0.003)	(0.005)	(0.003)
Education	0.049	-0.005	-0.002	-0.023	0.019
	(0.031)	(0.014)	(0.017)	(0.023)	(0.020)
Observations	4,884	4,884	4,884	4,884	4,884

Notes: We report only the parameter estimates and the standard deviation attached to the distribution factors. The variable Education refers to the education ratio computed as (1 + Edu_m) / (1 + Edu_f). The complete set of parameter estimates is available upon request. *** p<0.01, ** p<0.05, * p<0.1.

Table A4: Test of efficiency: "all or nothing" condition, younger heads

Budget shares	w1	w2	w3	w4	w5
Distribution factors:			*1998*		
Treatment	0.028**	-0.002	-0.007	-0.024**	-0.005
	(0.013)	(0.007)	(0.006)	(0.012)	(0.008)
Network	-0.008	-0.003	0.007	0.013*	0.008
	(0.009)	(0.005)	(0.005)	(0.007)	(0.005)
Education	0.040	-0.004	0.017	0.003	0.056**
	(0.032)	(0.021)	(0.020)	(0.027)	(0.023)
Observations	2,689	2,689	2,689	2,689	2,689
Distribution factors:			*1999*		
Treatment	-0.048***	-0.018**	0.022***	-0.001	-0.044***
	(0.012)	(0.007)	(0.005)	(0.009)	(0.008)
Network	0.009	0.008*	-0.002	0.007	0.021***
	(0.009)	(0.004)	(0.005)	(0.007)	(0.005)
Education	0.158***	0.015	-0.023	-0.091***	0.059***
	(0.034)	(0.016)	(0.018)	(0.030)	(0.022)
Observations	2,484	2,484	2,484	2,484	2,484
Distribution factors:			*Pooled*		
Treatment	0.002	-0.006	0.003	-0.016**	-0.020***
	(0.008)	(0.005)	(0.005)	(0.006)	(0.005)
Network	0.000	0.005	0.004	0.005	0.014***
	(0.006)	(0.003)	(0.003)	(0.005)	(0.004)
Education	0.051**	-0.006	0.022	-0.007	0.059***
	(0.026)	(0.012)	(0.014)	(0.021)	(0.016)
Observations	5,173	5,173	5,173	5,173	5,173

Notes: We report only the parameter estimates and the standard deviation attached to the distribution factors. The variable Education refers to the education ratio computed as (1 + Edu_m) / (1 + Edu_f). The complete set of parameter estimates is available upon request. *** p<0.01, ** p<0.05, * p<0.1.