Asymmetric Information in Households, with an Application to Bangladesh

Alexandre Fon

Abstract

This paper studies the effect of asymmetric information about income on household decisions, resource sharing, and welfare. It proceeds in four steps. In the first step, I develop a theoretical model that accounts for the possible existence of asymmetric information. The model predicts that households will partly mitigate the welfare cost of asymmetric information by incentivizing the wage earner to provide information about his or her true income. These incentives are provided by making the consumption share increase with reported income: the wage earner's consumption share is high when reporting a high income and low when reporting a low income. Second, I derive a new identification result for this model. The result states that individual welfare functions can be recovered with at least one assignable good - a good that is observed at the individual level rather than the household level. The result also shows that the effect of asymmetric information on individual welfare can be identified separately from other variables impacting welfare. Third, I estimate the model using a survey of Bangladeshi day laborers. The estimation confirms the predictions of the model, providing evidence that the households in the data are affected by asymmetric information. Finally, I conduct three counterfactual analyses to document how asymmetric information interacts with policies and compute the willingness to pay in each case.

1 Introduction

Analyzing the effect of policies on households requires considering the preferences and behavior of individual household members. The collective model (Chiappori 1988, 1992) is an effective and widely used way of doing precisely that. The collective model characterizes the household as a group of individuals, with possibly different preferences, making joint and efficient decisions. It has been used to analyze a large number of important economic questions, ranging from the optimal design of taxation systems and social safety nets, to poverty reduction in developing countries, to the education and labor force participation choices of women. To give a few examples, Low, Meghir, Pistaferri, Voena (2018) use the collective model to analyze PRWORA a major reform to the welfare system in the United States; Gayle and Shephard (2019) study optimal income taxation and its relationship with family structure; Attanasio and Lechene (2014) study PROGRESA, a conditional cash transfer to poor households in Mexico; Dunbar, Lewbel and Pendakur (2013, 2019) show that standard poverty indices can underestimate child poverty; Voena (2015) and Chiappori, Fortin and Lacroix (2002) show the link between marital contracts and female labor participation; Bronson (2015) uses the collective model to explain the increased college attendance of women since the 1970s. These examples highlight the prevalence and importance of the collective model and, as a consequence, how crucial it is to carefully examine the assumptions underlying the model.

An important assumption underlying the collective model is that household members have perfect information about each other's income. However, there is evidence that, for households in developing countries, asymmetric information is a source of inefficiency in the household. Field experiments, such as the one conducted by Ashraf (2009) in the Philippines, document that household members spend cash transfers differently depending on whether or not their spouse knows about the income. Experiments have found similar results in many developing countries (e.g., Castilla and Walker (2013), Hoel (2014), Ambler (2015), and Boltz, Marazyan, Villar (2019)). There is also some evidence that households in developed countries are affected. In a recent poll conducted in the US, 36% of respondents admitted to spending more than their spouse would be comfortable with, without their knowledge (YouGov 2020). This evidence indicates that if the standard collective model is used to answer policy questions, in particular in developing countries, it may lead to misleading conclusions. More generally, the evidence indicates the need for a framework to better understand how asymmetric information affects households.

This paper studies the effect of asymmetric information about income on household decisions, resource sharing, and welfare. In order to do this, this paper proceeds in four steps. First, it develops a model of a household facing asymmetric information about the income of the wage earner. The purpose of the model is to provide a framework to predict the effect of asymmetric information on household outcomes and to conduct counterfactual analyses. Second, a new identification result, specific to this model, shows that individual welfare functions can be estimated from consumption data. Third, the model is estimated using a sample of Bangladeshi day laborers. The estimation provides evidence that the day laborers are affected by asymmetric information. Finally, I conduct two counterfactual analyses of policies and compute the willingness-to-pay of the household.

The model introduced in this paper is a model of a household where wage earners have short-term income shocks that are unobservable to other household members. It relaxes and nests the perfect information assumption of the collective household model (Chiappori 1992) in order to account for the evidence on asymmetric information provided by previous papers.

The model shows that households making efficient decisions can reduce the cost of asymmetric information by incentivizing the wage earner to report their true income. This incentive is provided by making the individual consumption share of the wage earner increasing in the reported income. More precisely, if the wage earner reports a low income, they will get a smaller share of household consumption and if they report a high income they will get a bigger share of household consumption. As a consequence, they have an incentive to reveal their income. This reduces the welfare cost of asymmetric information. However, there will still be some cost caused by risk not being shared efficiently because the individual consumption of the wage earner responds "too much" to individual short-term income shocks. Therefore, the main prediction that emerges from the model is that, under asymmetric information, the consumption share of the wage earner will respond more to short-term income shocks than under perfect information.

The main modeling choice underlying this prediction is that households make efficient decisions. The reason for modeling households in this way is that many papers test and fail to reject efficiency in developing countries. While these tests are designed with perfect information in mind, the tests are still valid under asymmetric information. The reason is that, under asymmetric information, income realizations act as a shift in the decision power, but allocations are ex-post Pareto efficient for every realization. Most relevant for the context of my empirical results, Bargain, Lacroix, Tiberti (2018) document that data from Bangladeshi households are consistent with efficient decision-making. Other examples of papers that fail to reject efficiency in developing countries include Attanasio and Lechene (2014) and Bobonis (2009). Altruism between household members and coordination through repeated interactions are the main reasons that explain why households behave efficiently.

In order to bring the model to the data, a new nonparametric identification result specific to this model shows how individual welfare functions can be recovered from consumption data. In the data, consumption is typically observed at the household rather than individual level. This is the main challenge to identification in this literature. However, I show that a single "assignable good" – a good where individual consumptions are observed – is enough to recover the individual welfare functions, regardless of how many goods the household is consuming. In contrast to the identification result for the standard collective model (Chiappori and Ekeland 2009), no distribution factor - a variable impacting the decision power of households - is needed. However, the identification of this paper's model requires observing both average income, which is common knowledge in the household, as well as some measure of short-term income variation that can be affected by asymmetric information. The identification result also allows the effect of asymmetric information to be separately identified from other factors impacting household members' consumption shares.

The identification result is then used to estimate the model using a sample of Bangladeshi day laborers and their households. This data set is particularly suitable for the analysis for two reasons. First, families with day laborers are more likely to be affected by asymmetric information. Indeed, day laborers' jobs are often informal and have high short-term income variation, making it difficult for other household members to verify the day laborer's daily income. Second, the dataset contains two crucial variables for estimation. The first is food consumption at the individual level, which will serve as an assignable good. The second is short-term income variation, which is less likely to be observed by the spouse in this setting.

The estimation provides evidence that the households in the data are affected by asymmetric information. Since the model in this paper nests the case with perfect information, I can test whether perfect information is rejected by the data. The estimation shows that short-term income shocks lead to an increase in the day laborer's individual consumption share. This implies that household members are not sharing risk efficiently and allows me to reject perfect information. In order to confirm the asymmetric information mechanism, I document that short-term income shocks that are likely to be less observable by the spouse have a larger effect on individual consumption. Specifically, I find that income shocks of day laborers that work outside their village have a bigger impact. I also test the alternative hypothesis that the variation in consumption share is driven by nutritional needs: day laborers that work more have higher income and might have higher nutritional needs, which would lead to an increased consumption share. I do this by dividing income variation between daily wage variation and days worked variation. I find that income shocks due to variation in wages have a larger impact than variation in days worked, which provides evidence that the results are not being driven exclusively by nutritional needs. Overall, the evidence supports the model and the specific informational mechanism.

The estimation of the model makes it possible to conduct counterfactuals. The first counterfactual analysis quantifies the extra welfare cost if households do not make efficient decisions and do not incentivize truthful income reporting. The goal of this counterfactual is to help interpret past experimental results on asymmetric information. Experiments are one-off situations, which household members might not have experience navigating. As a result, households in experiments might not be able to provide incentives to mitigate the effect of asymmetric information. Therefore, experimental results might overstate the effect of asymmetric information compared to when it is a repeated issue facing the household. The simulation documents that without incentives, the day laborer chooses to hide 44% of the unobserved income. As a result, there are large welfare losses for the spouse and children, compared to when incentives are provided. Therefore, experiments might be considerably overstating the cost of asymmetric information.

I then conduct two policy counterfactuals. First, I consider a guaranteed employment scheme. This policy is used in developing countries, such as Bangladesh and India, to help poor workers through periods of low employment. The daily wage of participants is constant and fixed nationally, and therefore is not likely to be affected by asymmetric information. A policy simulation reveals that households are willing to pay 0.3% of total yearly household income to remove asymmetric information through a guaranteed employment scheme for a one year period. This indicates that, while asymmetric information does affect households, the welfare cost is quite small. Households are able to limit the welfare effects of asymmetric information by providing incentives to report income truthfully. Second, I consider a tax on a good that can be consumed with hidden income. I show that a 50% tax on such a good reduces the "extra" variance in consumption share due to asymmetric information from 26% to 16% of the total variance in consumption share.

This paper contributes to three main strands of literature. First, this paper contributes to the literature on cooperative models of intra-household allocations. In particular, it extends the standard collective model of the household which was first introduced by Chiappori (1988, 1992) and which is comprehensively reviewed by Chiappori and Mazzocco (2017). The extension in this paper is to allow for asymmetric information about income within the household. In addition, identification results from this literature (Chiappori and Ekeland 2009; Blundell, Chiappori, Meghir 2005) are extended to the model in this paper. More broadly, this paper contributes and builds on the literature on estimating collective models. The empirical specification used to estimate the model is closest to the one in Cherchye, De

Rock, Vermeulen's (2012) paper. Other papers in that have structurally estimated the collective model include Browning, Bourguignon, Chiappori and Lechene (1994); Dunbar, Lewbel, and Pendakur (2013, 2019); Mazzocco, Ruiz, and Yamaguchi (2014).

Second, this paper contributes to the literature on risk-sharing within the household. This paper provides evidence that households do not share risk efficiently. This is consistent with the results of Dercon and Krishnan (2000) and Duflo and Udry (2004) that find that poor households do not share risk efficiently in Ethiopia and the Ivory Coast respectively. My paper highlights the importance of asymmetric information in explaining this surprising result. Dubois and Ligon's (2002) paper is closest to this paper. They also document that consumption shares respond to short term income shocks and test whether asymmetric information is the cause. They conclude that asymmetric information is partly responsible for the observed variation in consumption shares. This paper differs along several dimensions. First, the source of asymmetric information Dubois and Ligon highlight concerns the effort of household members when working, while this paper focuses on asymmetric information about income. As a result, the testable implications that emerge from the two models are different. Second, this paper not only tests for asymmetric information, but also identifies and estimates the whole model, which then allows counterfactual analyses to be conducted. Mazzocco (2007) proposes a model with imperfect insurance within the household due to limited commitment of household members to future outcomes: household members will renegotiate their share of resources if their outside option of leaving the household increases enough. While limited commitment can explain imperfect insurance for larger income shocks, short-term income shocks are unlikely to significantly impact the outside option. Therefore, the asymmetric information mechanism proposed in this paper is likely to be the more relevant mechanism for imperfect insurance of short-term income shocks.

Third, many papers document the impact of asymmetric information about income on household behavior in experimental settings. For instance, Ashraf (2009) shows that individuals will spend windfall income differently depending on whether their spouse knows about the income. In this situation, one of the spouses takes advantage of the asymmetric information to consume goods that are more privately beneficial than if the income had been shared with the other spouse. Other papers also find evidence that asymmetric information matters: Castilla and Walker (2013); Hoel (2014); Ambler (2015). These papers provide very useful evidence that asymmetric information matters for household decisions. My analysis predicts that households that face asymmetric information repeatedly (like those with day laborers) can reduce the cost of asymmetric information through incentives. By contrast, experiments are typically one-offs and situations household members have no experience dealing with. As a result, household members are likely not able to coordinate to provide incentives in an experiment, as they would in a repeated setting outside of the lab. Therefore, the experimental evidence might be overestimating the impact of asymmetric information. An added benefit of the modeling approach I take is that I can estimate the model, which makes it possible to run counterfactuals.

The rest of the paper is organized as follows: the second part develops the model of the household. The third part derives an identification result for this model. The fourth part describes the data. The fifth part specifies a parametric empirical specification in order to estimate the model. The sixth part estimates the model and presents further results and counterfactuals. Finally, the seventh part concludes.

2 Model

2.1 Setting

This section describes the economic setting of a cooperative household that faces asymmetric information about income. The household is a two person household acting cooperatively. Agent 1, the wage earner draws random income $Y = \bar{y} + \tilde{y}$ where $\tilde{y}$ comes from a continuous distribution $F(\tilde{y})$ that is common knowledge. The income realization is not observed by agent 2, the home producer. Agent 1 has the possibility to consume a consumption good h_1 before revealing his income. The consumption of h_1 in this case is not observed by the household, allowing agent 1 to hide income and pretend to have had a lower income realization. The income revealed to the household is then used to consume two private goods (c and x) for each household member (c_1, x_1 for agent 1 and c_2, x_2 for agent 2) and a public good (X). The model can easily be generalized to a larger number of private and public goods. However, for the purpose of identification, which will be discussed later, there needs to be at least one good that is observed at the individual level in the data - c_i here. I will refer to c_i, x_i and X as "within-household" consumption. It is assumed that the preferences of agent i in the household can be represented by the following utility functions: $U^i(c_i, x_i, X) + v_i(h_i)$. In this model, agents have different preferences over h_i and the other private goods. This reflects the fact that hiding income limits consumption. In particular, consumption that would be observed by the spouse cannot be consumed with hidden income. Therefore, we can think that in general, consuming the hidden good has a "cost" reflected in the preferences, which comes from the fact that your consumption options are limited.

The presence of asymmetric information constrains the possible consumption allocations that can be reached in equilibrium. The allocations have to be incentive compatible for agent 1 − i.e. he is weakly better off by revealing his true income than by pretending to have a lower income and consuming the difference in incomes as hidden consumption. Subject to the incentive compatibility constraint, the outcomes are assumed to be Pareto efficient. The household's problem is the following[1]:

$$\max_{c_i,x_i,h_i,X,i=1,2\ \forall\tilde{y}} \mathbb{E}_{\tilde{y}}\ \mu(\bar{y},P,\pi)[U^1(c_1,x_1,X)+v_1(h_1)]+(1-\mu(\bar{y},P,\pi))[U^2(c_2,x_2,X)+v_2(h_2)]$$

$$\text{s.t.} \qquad c_1+c_2+px_1+px_2+\pi X+qh_1+qh_2=\bar{y}+\tilde{y} \qquad \forall\tilde{y}$$

$$U^1(c_1,x_1,X)+v_1(h_1)\geq U^1(c_1',x',X')+v_1(h_1'+\frac{\tilde{y}-\tilde{y}'}{q}) \quad \forall\tilde{y}'<\tilde{y}$$

In the above, the first constraint is the budget constraint and the second constraint is the IC constraint. The IC constraint says that agent 1 must weakly prefer the equilibrium allocation (on the left-hand side) to deviating by pretending to have a lower income draw and spending the difference in income on the outside good (on the right-hand side). p is the price of good x, π is the price of the public good and q the price of the hidden consumption good. μ the decision power of agent 1. It can depend arbitrarily on the average income $\bar{y}$, the vector of private prices $P=(p,q)$ and π. Here it is assumed that agent 1 is egoistic (the IC constraint depends only on the preferences of agent 1). However, it is easy to add "separable" altruism in this model: agent 1's utility is a weighted sum of his and agent 2's egoistic utility functions. The predictions of the model would be unchanged by adding this specific form of altruism.

There is a slight abuse of notation in the problem above. Technically, each decision variable is a function of $\tilde{y}$ since we have to solve the problem for each realization of income. We suppress the function for ease of notation but it is important to remember that we are solving for a function for each of these decision variables. Note that if it weren't for the IC constraint the household's problems for any two realizations of $\tilde{y}$ would be independent. The IC constraint connects the problem for two realizations and makes it such that maximizing the household's expected utility is more difficult than simply maximizing the utility at every realization of $\tilde{y}$. Such a naive approach would not respect the IC constraint and therefore agent 1 would have an incentive to deviate.

[1]A solution to the following problem will be Pareto efficient. However, the converse is not necessarily true. The standard proof is based on the separating hyperplane theorem that applies to convex sets. Here the set of possible allocations - that are both feasible and IC - is not necessarily convex because of the IC constraint

2.2 IC constraint

The IC constraint in this specific form is a complicated object. Therefore, we can manipulate it to make it easier to deal with. Let us denote $IC(\tilde{y}, \tilde{y}')$ the IC constraint for a specific $\tilde{y}$ and $\tilde{y}'$ with $\tilde{y} > \tilde{y}'$. Under the assumption that $v(.)$ is weakly concave, it is possible to show that for any $\tilde{y} > \tilde{y}' > \tilde{y}''$, $IC(\tilde{y}, \tilde{y}')$ and $IC(\tilde{y}', \tilde{y}'')$ imply $IC(\tilde{y}, \tilde{y}'')$. Therefore, if the IC constraint holds locally, it holds globally (if agent 1 does not have an incentive to pretend to have income slightly below his true income he does not have an an incentive to pretend to have any other income). If the distribution of income is discrete, any realization of income $\tilde{y}$ only appears in two binding IC constraints: with the value just above $\tilde{y}$ and with the value just below $\tilde{y}$. Since here the distribution of income is continuous the IC constraint will constrain the derivative of agent 1's indirect utility. To show this let us rewrite the IC constraint above in terms of the utility level of within-household consumption when reporting income $\tilde{y}$, $\tilde{U}^1(\tilde{y}) \equiv U^1(c_1(\tilde{y}), x_1(\tilde{y}), X(\tilde{y}))$:

$$\tilde{U}^1(\tilde{y}) + v_1(h_1(\tilde{y})) \geq \tilde{U}^1(\tilde{y}) + v_1\big(h_1(\tilde{y}') + \frac{\tilde{y} - \tilde{y}'}{q}\big) \qquad \forall \quad \tilde{y} > \tilde{y}'$$

We can divide on both sides by $\tilde{y} - \tilde{y}'$. Rearranging, we get:

$$\frac{\tilde{U}^1(\tilde{y}) - \tilde{U}^1(\tilde{y}')}{\tilde{y} - \tilde{y}'} \geq -\frac{v_1(h_1(\tilde{y})) - v_1(h_1(\tilde{y}') + \frac{\tilde{y}-\tilde{y}'}{q})}{\tilde{y} - \tilde{y}'} \qquad \forall \quad \tilde{y} > \tilde{y}'$$

Taking the limit as $\tilde{y}'$ goes to $\tilde{y}$, we get

$$\tilde{U}'^1(\tilde{y}) \geq v_1'(h_1(\tilde{y}))\big(\frac{1}{q} - \frac{\partial h_1(\tilde{y})}{\partial \tilde{y}}\big)$$

We can rewrite this in an intuitive manner:

$$\tilde{U}'^1(\tilde{y}) + v_1'(h_1(\tilde{y}))\frac{\partial h_1(\tilde{y})}{\partial \tilde{y}} \geq \frac{1}{q} v_1'(h_1(\tilde{y})) \qquad (1)$$

This equation states that the marginal utility from revealing true income has to be greater or equal than the marginal utility from hiding income for agent 1. The marginal utility from revealing true income is the sum of the marginal utility from within-household consumption and the marginal utility from the change in hidden consumption in the equilibrium allocation. Meanwhile, the marginal utility from hiding income is simply the increased utility from consuming only the hidden consumption good with the hidden income.

Integrating equation (1) on both sides with respect to $\tilde{y}$ between some arbitrary $\tilde{y}$ and $\tilde{y}_0$ (the smallest possible value) we get the following, where $U_0^1 = U^1(c_1(\tilde{y}_0), x_1(\tilde{y}_0), X(\tilde{y}_0)) + v_1(h_1(\tilde{y}_0))$ is agent 1's utility when he truthfully reports the lowest income realization $\tilde{y}_0$:

$$U^1(c_1(\tilde{y}), x_1(\tilde{y}), X(\tilde{y})) + v_1(h_1(\tilde{y})) - U_0^1 \geq \int_{\tilde{y}_0}^{y} v_1'(h_1(t))\mathrm{d}t \tag{2}$$

Note that in the special case where $v_1(h_1)$ is linear, the RHS of equation (2) becomes a linear function of $\tilde{y} - \tilde{y}_0$. In that case, the IC constraint will restrict agent 1's utility to increase linearly in income. The more general case in equation (3) is similar. The main difference is that the slope of agent 1's utility is not constant but rather decreases with $h_1(\tilde{y})$ (under the assumption that $v_1(.)$ is concave).

For the remainder, I will assume that the IC constraint binds for all $\tilde{y}$. Therefore, I will treat equation (3) as an equality.

2.3 Solution

We can therefore rewrite the household's problem as:

$$\begin{aligned}
\max_{c_i, x_i, h_i, X,\ \forall \tilde{y}} \quad & \mathbb{E}_{\tilde{y}} \ \mu(\bar{y}, P, \pi)[U^1(c_1, x_1, X) + v_1(h_1)] + (1 - \mu(\bar{y}, P, \pi))[U^2(c_2, x_2, X) + v_2(h_2)] \\
\text{s.t.} \quad & c_1 + c_2 + px_1 + px_2 + \pi X + qh_1 + qh_2 = \bar{y} + \tilde{y} \qquad \forall \tilde{y} \\
& U^1(c_1, x_1, X) + v_1(h_1) \geq U_0^1 + \frac{1}{q} \int_{y^0}^{\tilde{y}} v_1'(h_1(t))\mathrm{d}t
\end{aligned}$$

We can take FOCs to solve this problem. λ_1 and λ_2 are the Lagrange multipliers associated with the budget constraint and the IC-constraint respectively. I will denote by U_k^i the derivative of the utility function of agent i with respect to argument k. I also will not write the decision variables as functions of y explicitly for ease of notation

$$(\mu + \lambda_2)U_1^1(c_1, x_1, X) = \lambda_1 \tag{c_1}$$

$$(1-\mu)U_1^2(c_2, x_2, X) = \lambda_1 \tag{c_2}$$

$$(\mu+\lambda_2)U_2^1(c_1, x_1, X) = \lambda_1 \tag{x_1}$$

$$(1-\mu)U_2^2(c_2, x_2, X) = \lambda_1 \tag{x_2}$$

$$(\mu+\lambda_2)U_3^1(c_1, x_1, X) + (1-\mu)U_3^2(c_2, x_2, X) = \pi\lambda_1 \tag{X}$$

$$(\mu+\lambda_2)v_1'(h_1(y)) = q\lambda_1 + \lambda_2 \frac{1}{q} v_1'(h_1(y)) \frac{1}{h_1'(y)} \tag{h_1}$$

$$(1-\mu)v_2'(h_2) = q\lambda_1 \tag{h_2}$$

This set of equations defines the allocation for each income realization $\tilde{y}$. The first thing to note is that this set of equations is very similar to the set of equations we would have if there was no IC constraint. In fact, apart from the FOC for h_1, the other equations are exactly the same that we would get if we solved the household's problem with Pareto weight on agent 1 being $\mu + \lambda_2$ and on agent 2 being $1 - \mu$. Normalizing the Pareto weight to sum to 1 this would imply that the Pareto weight of agent 1 will be $\frac{\mu+\lambda_2}{1+\lambda_2}$ instead of μ. Therefore, for all goods apart from h_1 the IC constraints can be thought of as simply shifting the decision power of agent 1. This can be seen most clearly in the modified efficiency rule:

$$\frac{\mu+\lambda_2}{1-\mu} = \frac{U_1^2(c_2, x_2, X)}{U_1^1(c_1, x_1, X)}$$

Under perfect information, we would have $\lambda_2 = 0$ for all income realizations. However, with asymmetric information, the ratio of marginal utilities will deviate from the ratio of Pareto weights.

Looking at equation (2), we can see that with a concave $v(.)$, changing consumption of hidden income can make the IC-constraint slacker or tighter. Therefore, it is to be expected that the optimal consumption of h_1 would take this effect into account. More precisely, what this tells us is that the marginal benefit of consuming h_1 has to equal to marginal cost. This marginal cost has the usual BC term $q\lambda_1$ and an additional term. When λ_2 is negative, this cost is also negative, therefore the wage earner will consume more h_1. The inverse happens when λ_2 is positive.

An important object in this set of equations is λ_2. $\lambda_2/(1+\lambda_2)$ is the additional decision power of agent 1 for any given income realization. Solving for λ_2, we find that:

$$\lambda_2 = \frac{(1-\mu)U_1^2(c_2, x_2, X) - \mu U_1^1(c_1, x_1, X)}{U_1^1(c_1, x_1, X)} = -\mu + (1-\mu)\frac{U_1^2(c_2, x_2, X)}{U_1^1(c_1, x_1, X)} \tag{3}$$

λ_2, the multiplier on the IC constraint, has an intuitive form. We have seen that the choice of c_1, x_1 and X is efficient for any given realization of y, with only the decision power shifting. Therefore, λ_2 is simply by how much the ratio of marginal utilities deviates from the standard efficiency condition. Note that λ_2 can be positive or negative depending on which marginal utility is "too high".

For a given U_0, the set of FOCs above, along with the budget constraint and the IC constraint define the optimal allocation. Therefore, to completely characterize the optimal allocation solving for U_0 is the only remaining step.

Taking the FOC for U_0 we get:

$$\mathbb{E}_y[-\lambda_2] = 0$$

Plugging in the expression for λ_2 given in equation (3) and rearranging, we get:

$$\mathbb{E}_y\left[\frac{U_1^2(c_2, x_2, X)}{U_1^1 c_1, x_1, X}\right] = \frac{\mu}{1-\mu}$$

Therefore we find that even though the efficiency rule does not hold at each $\tilde{y}$, U_0^1 is chosen such that it holds on average.

2.4 Two-Stage Decision Process

A helpful tool to think about identification is to break down the household's decision in a two-stage decision process. It allows us to define a *conditional sharing-rule* (this idea is standard in collective household models for example Blundell, Chiappori, Meghir (2005)). The sharing-rule summarizes how household members share ressources and therefore selects one particular outcome out of the set of contrained efficient outcomes. In that capacity, it serves the same purpose as the Pareto weights but with the added benefits of being more clearly interpretable and not depending on the particular cardinalization of utilities.

The two stages will be the following. In the first stage, the household jointly choose U_0^1, h_i, X and how to split the residual income between them. In the second stage, each household member freely chooses how to allocate this income between the two private goods c_i and x_i. More formally, denote by $c_i(P,\pi,\bar{y},\tilde{y})$, $x_i(P,\pi,\bar{y},\tilde{y})$, $h_i(P,\pi,\bar{y},\tilde{y})$, $X(P,\pi,\bar{y},\tilde{y})$ the solution to the household's problem. Then define the sharing-rule: $r(P,\pi,\bar{y},\tilde{y}) = c_1(P,\pi,\bar{y},\tilde{y}) + px_1(P,\pi,\bar{y},\tilde{y})$. Then we will have:

$$r^1(P,\pi,\bar{y},\tilde{y}) \equiv r(P,\pi,\bar{y},\tilde{y})$$

$$r^2(P,\pi,\bar{y},\tilde{y}) \equiv \bar{y} + \tilde{y} - \pi X(P,\pi,\bar{y},\tilde{y}) - q(h_1(P,\pi,\bar{y},\tilde{y}) + h_2(P,\pi,\bar{y},\tilde{y})) - r(P,\pi,\bar{y},\tilde{y})$$

from the budget constraint. The r^i functions simply tell us how much of the residual income after consuming the public good and the hideable good is given to each household member i.

Proposition 1. Let $c_i^* = c_i(P,\pi,\bar{y},\tilde{y})$, $x_i^* = x_i(P,\pi,\bar{y},\tilde{y})$, $h_i^* = h_i(P,\pi,\bar{y},\tilde{y})$, $X^* = X(P,\pi,\bar{y},\tilde{y})$ be the solution to the household's problem defined in (1). Then $c_i(P,\pi,\bar{y},\tilde{y})$, $x_i(P,\pi,\bar{y},\tilde{y})$ solves:

$$\begin{aligned} \max_{c_i,x_i} \quad & U_i(c_i,x_i,X^*) \\ \text{s.t.} \quad & c_i + px_i = r^i(P,\pi,\bar{y},\tilde{y}) \end{aligned} \tag{4}$$

Proof for agent 2. For agent 2 this is a very simple proof by contradiction. Suppose that the proposition does not hold for agent 2. Then there must be a c_2', x_2' that respects the budget constraint such that $U_2(c_2',x_2',X^*) > U_2(c_2^*,x_2^*,X^*)$. However, if that were the case then the maximand in problem (1) could be increased by replacing c_2^* and x_2^* by c_2' and x_2' while still satsifying all the constraints, which is a contradiction. □

For agent 1, we just need to be a little more careful because the utility for a given $\tilde{y}$ is constrained by the IC constraint. In particular, an increase in the utility of agent 1 at a given income realization might not be feasible in the household's problem because of the IC constraint. However, the separability[2] of outside consumption h_1 means that only the within-household utility level matters (U_1^*) for the IC constraint and not which goods are being consumed. Therefore, agent 1 will be optimizing in the second stage.

Proof for agent 1. Suppose that the proposition does not hold for agent 1. Then there must be a c_1', x_1' that respects the budget constraint such that $U_1(c_1', x_1', X^*) > U_1(c_1^*, x_1^*, X^*)$. Then, since utilities are increasing in the various arguments, there must a c_1'', x_1'' such that $U_1(c_1'', x_1'', X^*) = U_i(c_1^*, x_1^*, X^*)$ but $c_1'' + px_1'' < r_1^*$. Then, we have left-over income that we can give to agent 2. For example, the bundle (c_1'', x_1'', $c_2'' = c_2^* + (r_1^* - (c_1'' + px_1''))$, x_2^*, X^*, h_1^*, h_2^*) will increase the maximand from the household's problem. In addition, the budget constraint holds by construction since the extra consumption for agent 2 corresponds precisely to the left over income from agent 1. Finally, the IC constraint will hold since it only depends on c_1 and x_1 through the within-household utility level of agent 1 (U_1^*) which hasn't changed. Therefore, the maximand of the household's problem can be increased while still respecting all constraints, which is a contradiction. □

2.5 Indirect Utilities

We are going to define two relevant indirect utility concepts here. Again, this follows Blundell, Chiappori and Meghir (2005) and Chiappori and Ekeland (2009). The first indirect utility is called the conditional individual indirect utility. It is the maximized value of program number (2) for some arbitrary values of r^i and X. Therefore it is the value of agent i's maximized utility as function of p, r^i and X and it is denoted $V^i(p, r^i, X)$. It is called "individual" because it only depends on agent i's preferences. It is called "conditional" because it depends on X.

Obviously, if we know individual preferences $V^i(p, r^i, X)$ and the decision process $r^i(P, \pi, \bar{y}, \tilde{y})$ we know everything there is to know about the household. However, as we shall see it is not possible to identify these two functions separately.

Instead of conditional individual indirect utilities, the identification will aim to recover the conditional collective indirect utilities. This indirect utility is called "collective" because it combines individual preferences and the decision process. It is still called conditional because it still depends on X.

[2]Here we have strong separability but weak separability would be enough for the identification proof to hold.

To get this function, two changes of variables are required. The first change of variable is specific to the setting with asymmetric information. One specificity of this setting is that the outside good h_1 has to be chosen in the first stage because it affects directly the IC constraint and not only through agent 1's within-household utility function[3]. Then, notice that the h_i enter the definition of the function ρ^2 which will complicate the identification. Therefore, we can make the following change of variables. Denote $y = \bar{y} + \tilde{y} - \pi X(P, \pi, \bar{y}, \tilde{y}) - q(h_1(P, \pi, \bar{y}, \tilde{y}) + h_2(P, \pi, \bar{y}, \tilde{y}))$ the residual income in stage 1 after consuming the public good and the outside good. Then, as long as the partial of residual income with respect to $\bar{y}$ is not 0 [4] we can locally apply the implicit function theorem to express $\bar{y}$ as some function ϕ of the other exogenous variables and residual income. We can then plug ϕ into the sharing-rule. We then define the new sharing-rule function $\tilde{r}(P, \pi, y, \tilde{y}) = r(P, \pi, \phi(P, \pi, y, \tilde{y}), \tilde{y})$ and we then have $\tilde{r}^1(P, \pi, y, \tilde{y}) = \tilde{r}(P, \pi, y, \tilde{y})$ and $\tilde{r}^2(P, \pi, y, \tilde{y}) = y - \tilde{r}(P, \pi, y, \tilde{y})$. Since residual income is observed directly this won't complicate identification.

The second change of variable is the standard one in collective models with public consumption. We will vary the price of the public good π such that X is kept constant. Take a neighbourhood of some point $(P, \pi, y, \tilde{y})$ where $\dfrac{\partial X(P, \pi, y, \tilde{y})}{\partial \pi} \neq 0$. By the implicit function theorem we can use the equation $X(P, \pi, y, \tilde{y}) = X$ to get π as some function ψ of all the other exogenous variables and X. Then we can plug this function into the sharing-rule $\tilde{r}$ to get the sharing-rule as a function of public consumption X.

Then we can define the conditional sharing rule:

$$\rho(P, y, \tilde{y}, X) = \tilde{r}(P, \psi(P, y, \tilde{y}, X), y, \tilde{y}) \tag{5}$$

and the conditional collective indirect utility:

$$W^i(P, y, \tilde{y}, X) = V^i(p, \rho^i(P, y, \tilde{y}, X), X) \tag{6}$$

Three things are important to note here. First, we will have that $\rho^1 \equiv \rho(P, y, \tilde{y}, X)$ and $\rho^2(P, y, \tilde{y}, X) \equiv y - \rho(P, y, \tilde{y}, X)$. Second, the collective indirect utility function is the one that is relevant for welfare analysis because it summarizes the utility outcome for agent i of any policy, taking into account the redistribution within the household. Third, since $X(P, \pi, y, \tilde{y})$ and $\psi(P, y, \tilde{y}, X)$ are known functions, it is straightforward to do the change of variable described above in either direction. In particular, if we identify the conditional

[3] h_1 has to be chosen in the first stage because the choice of h_1 can make the IC constraint tighter or slacker, affecting the ressource split. Therefore, it is similar to a public good because both household members care about the consumption of that good.

[4] It would only be 0 if all the extra income would be used to purchase h_1, h_2 and X which seems unlikely

collective indirect utility it is easy to get the unconditional one by replacing X by the function $X(P, \pi, y, \tilde{y})$.

2.6 First Stage of the Household Problem

Now that we've defined the indirect utilities from the second stage of the household decision process we have the tools to write explicitly the first stage. The individual indirect utility is the relevant concept when it comes to writing the first-stage in the two-stage representation of the household's problem. The household decisions in the first stage are made knowing that in the second stage each household member will maximize their utility given the choice of r, X and the value of p. Therefore, the household knows that the choice of X and r will lead to individual utilities $V^i(p, r^i, X)$.

This allows us to write the first stage of the household problem:

$$\begin{aligned} \max_{r^i, h_i, X, \; \forall \tilde{y}} \quad & \mathbb{E}_{\tilde{y}} \;\; \mu(\bar{y}, P, \pi)[V^1(p, r^1, X) + v_1(h_1)] + (1 - \mu(\bar{y}, P, \pi))[V^2(p, r^2, X) + v_2(h_2)] \\ \text{s.t.} \quad & r^1 + r^2 + \pi X + q h_1 + q h_2 = \bar{y} + \tilde{y} \qquad \forall \tilde{y} \\ & V^1(p, r^1, X) + v_1(h_1) \geq V_0^1 + \frac{1}{q} \int_{y^0}^{\tilde{y}} v_1'(h_1(t)) \mathrm{d}t \end{aligned} \tag{7}$$

3 Identification

As is common in collective models, a main challenge of identification is that the individual consumption of private goods is not observed separately for all the goods. In particular, in this model, we will assume that we can observe the individual consumption of one of the private goods (c_i with loss of generality) but not of the other (x_i). Rather, for consumption good x we only observe the household aggregate demand $x = x_1 + x_2$. Following the two changes of variables in the previous section the situation is now very similar to Chiappori and Ekeland (2009). In particular, $\tilde{y}$ can play the role of a distribution factor. Therefore, one assignable good is sufficient to identify the collective indirect utilities.[5] More precisely, the knowledge of c_1, c_2, $x = x_1 + x_2$, X as functions of p, π, y and $\tilde{y}$ is sufficient to recover the welfare relevant structure. Details are given below.

[5] In the terminology of the collective model, an assignable good is a good where the consumption of each member is observed separately.

3.1 Identification of the Sharing-Rule

Denote $A = \dfrac{\partial c_1(P,y,\tilde{y},X)}{\partial y} \Big/ \dfrac{\partial c_1(P,y,\tilde{y},X)}{\partial \tilde{y}}$, $B = \dfrac{\partial c_2(P,y,\tilde{y},X)}{\partial y} \Big/ \dfrac{\partial c_2(P,y,\tilde{y},X)}{\partial \tilde{y}}$,

$C = \dfrac{\partial c_1(P,y,\tilde{y},X)}{\partial y} \Big/ \dfrac{\partial c_1(P,y,\tilde{y},X)}{\partial q}$ and $D = \dfrac{\partial c_2(P,y,\tilde{y},X)}{\partial y} \Big/ \dfrac{\partial c_2(P,y,\tilde{y},X)}{\partial q}$.

Proposition 2. If $A \neq B$ and $C \neq D$, then, the knowledge of the continuous, differentiable c_1, c_2, $x = x_1 + x_2$, X as functions of P, π, y and $\tilde{y}$ allows us to identify the sharing-rule up to an additive function of p and X. That is, two sharing-rules $\rho(P,y,\tilde{y},X)$ and $\hat{\rho}(P,y,\tilde{y},X)$ will be such that $\hat{\rho}(P,y,\tilde{y},X) = \rho(P,y,\tilde{y},X) + f(p,X)$

Proof. From the two stage formulation we have that $c_i(P,y,\tilde{y},X) = c_i^*(p,\rho^i(P,y,\tilde{y},X),X)$.

Recall that $\rho^1(P,y,\tilde{y},Q) = \rho(P,y,\tilde{y},X)$ and $\rho^2(P,y,\tilde{y},X) = y - \rho(P,y,\tilde{y},X)$. Then, from the two-stage formulation, we have that $A = \dfrac{\rho_y}{\rho_{\tilde{y}}}$ and $B = -\dfrac{1-\rho_y}{\rho_{\tilde{y}}}$ where ρ_y denotes the partial of the function $\rho = \rho^1$ with respect to y and similarly for $\tilde{y}$. A and B are known from the knowledge of the demand functions. Solving, we find $\rho_y = \frac{A}{A-B}$ and $\rho_{\tilde{y}} = \frac{1}{A-B}$. Similarly, from C and D we find that $\rho_y = \frac{C}{C-D}$ and $\rho_q = \frac{1}{C-D}$. Note that we are overidentified here since we find two independent expressions for ρ_y[6]. We recovered three of the partial derivatives of ρ. Therefore, if ρ and ρ' are two sharing-rules the difference will be of the form $f(p,X)$. Alternatively, we have identified ρ up to an additive function of p and X.

□

3.2 Identification of the Collective Indirect Utilities

Proposition 3. Under the same conditions as proposition 2, the conditional collective indirect utilities are identified up to an increasing transformation.

Proof. Corollary 8 and therefore proposition 7 from Chiappori and Ekeland (2009) apply here. Proposition 7 gives conditions under which the collective indirect utilities are identified up to an increasing transformation. These conditions hold here.

[6] q and $\tilde{y}$ enter this problem similarly to distribution factors in the collective model. This overidentification leads to a natural test of the model that is similar to the well known test of proportionality of distribution factors.

The first thing to note is that we have exactly the same structure as that paper regarding the within-household utilities W^i. The difference between the two settings is that this paper has asymmetric information. However, we have already seen that in equilibrium this will affect the allocations of within-household consumptions the same way a variation in the decision power would. In Chiappori and Ekeland, the decision power can vary arbitrarily with the exogenous variables. Therefore, the decision power varying in our setting is not an issue. In addition, thanks to our change of variables the choice of h_i does not affect the second stage. Finally, the exogenous variables that do not have an equivalent in their paper (q and $\tilde{y}$) enter the the second stage only through the sharing rule. Therefore, they are isomorphic to *distribution factors* in their setting.

Since the two settings are equivalent regarding within-household consumption, the same conditions are necessary to apply corollary 8. More precisely, applying that result requires the collective indirect utilities W^i to be "separable through ρ^i". This simply means that we need:

$$\frac{\partial W^i(P,y,\tilde{y},X)}{\partial y}\frac{\partial \rho^i}{\partial \tilde{y}} = \frac{\partial W^i(P,y,\tilde{y},X)}{\partial \tilde{y}}\frac{\partial \rho^i}{\partial y}$$

To see why this holds in this setting, take derivatives of equation (6) with respect to y and $\tilde{y}$. Intuitively, this separability simply means that the variables y and $\tilde{y}$ enter each agent's indirect utility through the same function – the sharing-rule. Since W^i is separable through ρ^i, we can apply corollary 8 and as a consequence proposition 7 of Chiappori and Ekeland. Therefore the collective indirect utilities are identified. □

This next part will discuss more informally the identification idea. Many of the arguments are adpated from Ekeland and Chiappori (2009) and Bluncdell, Chiappori, and Meghir (2005).

Fix (p, X) for this part of the analysis. Recall that the vector P is the vector of the prices for the private goods (p, q). We have the following:

$$\frac{\partial W^i(P,y,\tilde{y},X)}{\partial \tilde{y}}\Big/\frac{\partial W^i(P,y,\tilde{y},X)}{\partial y} = \frac{\rho^i_{\tilde{y}}}{\rho^i_y} \tag{8}$$

$$\frac{\partial W^i(P,y,\tilde{y},X)}{\partial q}\Big/\frac{\partial W^i(P,y,\tilde{y},X)}{\partial y} = \frac{\partial \rho^i}{\partial q}\Big/\frac{\partial \rho^i}{\partial y} \tag{9}$$

The right-hand side of the equalities above are known. Therefore, we've identified the ratios of partial derivatives of the collective indirect utility. This is equivalent to identifying the function $W^i(.)$ up to an increasing transformation at a fixed (p, X). Given that this function describes preferences it is to be expected that we would only be able to identify it up to an increasing transformation. See the appendix for a proof that knowing the ratios of partial derivatives is equivalent to knowing the function up to an increasing transformation.

Note that the collective indirect utility can still vary arbitrarily with p and X. Therefore, the knowledge of $\rho_{\tilde{y}}, \rho_y, \rho_q$ has allowed us to identify the collective indirect utilities W^i up to an increasing transformation that can depend on p and X. Explicitely, at this stage, we know that the relationship between two collective indirect utilities $\tilde{W}^i$ and W^i that respect equations (8) and (9) must be the following, for some function F^i increasing in it's first argument:

$$W^i(P, y, \tilde{y}, X) = F^i(\tilde{W}^i(P, y, \tilde{y}, X), p, X) \tag{10}$$

At this stage, we do not know anything about $\frac{\partial W^i(P,y,\tilde{y},X)}{\partial p}$ and $\frac{\partial W^i(P,y,\tilde{y},X)}{\partial X}$. These two partial derivatives are not constrained by equations (8) and (9). The goal of this next part will be precisely to see what we can learn about these two partial derivatives. Towards this goal let us pick one arbitrary function $\tilde{W}^i$ that respects equations (8) and (9). Obviously, this will not be an actual collective indirect utility because at this stage we do not know anything about how W^i varies with p and X. The question then becomes: for a given $\tilde{W}^i$, can we recover a unique function F_i (up to an increasing transformation) such that the true collective indirect utility W^i is given by $W^i(P, y, \tilde{y}, X) = F^i(\tilde{W}^i(P, y, \tilde{y}, X), p, X)$? The rest of this section answers positively and describes the steps.

Let us first focus on $\dfrac{\partial W^i(P, y, \tilde{y}, X)}{\partial p}$. By the enveloppe theorem applied to program (4) we find:

$$\frac{\partial V^i(p, r^i, X)}{\partial p} \Big/ \frac{\partial V^i(p, r^i, X)}{\partial r^i} = x_i$$

Now, using the relationship between W^i and V^i described in equation (6) we find:

$$\frac{\partial W^i(P, y, \tilde{y}, X)}{\partial p} \Big/ \frac{\partial W^i(P, y, \tilde{y}, X)}{\partial \tilde{y}} = \frac{x_i + \rho_p^i}{\rho_{\tilde{y}}^i}$$

Then, since $\rho^1 = \rho$ and $\rho^2 = y - \rho$ we have: $\rho_p = \rho_p^1 = -\rho_p^2$ and $\rho_{\tilde{y}} = \rho_{\tilde{y}}^1 = -\rho_{\tilde{y}}^2$. Therefore we get the following where the right hand side is observable and known:

$$\frac{\partial W^1(P, y, \tilde{y}, X)}{\partial p} \Big/ \frac{\partial W^1(P, y, \tilde{y}, X)}{\partial \tilde{y}} - \frac{\partial W^2(P, y, \tilde{y}, X)}{\partial p} \Big/ \frac{\partial W^2(P, y, \tilde{y}, X)}{\partial \tilde{y}} = \frac{x_1 + x_2}{\rho_{\tilde{y}}} = \frac{x(P, y, \tilde{y}, X)}{\rho_{\tilde{y}}} \tag{11}$$

Now we want to identify the functions F^i. From equation (10) we can express the partial derivatives of W^i as functions of F^i and $\tilde{W}^i$ and replace them into (11). Recall that

$\tilde{W}_i$ are functions that we have chosen and are therefore known. We get the following:

$$\Big(\frac{\partial F^1}{\partial \tilde{W}^1}\frac{\partial \tilde{W}^1}{\partial p}+\frac{\partial F^1}{\partial p}\Big)\Big/\frac{\partial F^1}{\partial \tilde{W}^1}\frac{\partial \tilde{W}^1}{\partial \tilde{y}}-\big(\frac{\partial F^2}{\partial \tilde{W}^2}\frac{\partial \tilde{W}^2}{\partial p}+\frac{\partial F^2}{\partial p}\Big)\Big/\frac{\partial F^2}{\partial \tilde{W}^2}\frac{\partial \tilde{W}_2}{\partial \tilde{y}}=\frac{x(P,y,\tilde{y},X)}{\rho_{\tilde{y}}}$$

Rearranging we get the following, where the right hand side is known:

$$\frac{\partial F^1}{\partial p}\Big/\frac{\partial F^1}{\partial \tilde{W}^1}\Big(\frac{\partial \tilde{W}^1}{\partial \tilde{y}}\Big)^{-1}-\frac{\partial F^2}{\partial p}\Big/\frac{\partial F^2}{\partial \tilde{W}^2}\Big(\frac{\partial \tilde{W}^2}{\partial \tilde{y}}\Big)^{-1}=\frac{x}{\rho_{\tilde{y}}}-\frac{\partial \tilde{W}^1}{\partial p}\Big/\frac{\partial \tilde{W}^1}{\partial \tilde{y}}+\frac{\partial \tilde{W}^1}{\partial p}\Big/\frac{\partial \tilde{W}^2}{\partial \tilde{y}} \quad (12)$$

Since utilities can only be identified up to an increasing transformation, the best we can hope for is to identify the ratios $\gamma^i(\tilde{W}^i,p,X)\equiv\frac{\partial F^i}{\partial p}\Big/\frac{\partial F^i}{\partial W^i}$. Then we must show that the solution to (12) in terms of the ratios is unique. Intuitively, since γ^i depends only on 3 variables but $\frac{\partial W^i}{\partial \tilde{y}}$ depends on $(P,y,\tilde{y},X)$ we can expect there to be only one solution. In practice, this will be the case almost always. This is usually described as "generic" identification in the collective model literature. See Blundell, Chiappori and Meghir (2005) for a more precise discussion of genericity.

Let us give some intuition of why the solution to (12) is unique. Suppose, by way of contradiction, that there are two solutions (γ^1,γ^2) and $(\gamma^{1'},\gamma^{2'})$. Denote $\delta^i=\gamma^i-\gamma^{i'}$. Then we can take the difference of (12) for the two solutions to get:

$$\delta^1(\tilde{W}^1,p,X)\Big(\frac{\partial \tilde{W}^1}{\partial \tilde{y}}\Big)^{-1}-\delta^2(\tilde{W}^2,p,X)\Big(\frac{\partial \tilde{W}^2}{\partial \tilde{y}}\Big)^{-1}=0 \quad (13)$$

Then, for any point at which $\delta^i(\tilde{W}^i,p,X)\neq 0$ then $\delta^j(\tilde{W}^i,p,X)\neq 0$ for $i\neq j$ and we can write (13) as:

$$\log\delta^1(\tilde{W}^1,p,X)-\log\delta^2(\tilde{W}^2,p,X)=\log\Big(\frac{\partial \tilde{W}^1}{\partial \tilde{y}}\Big/\frac{\partial \tilde{W}^2}{\partial \tilde{y}}\Big) \quad (14)$$

This implies that the right hand side must be equal to the sum of a function of $(\tilde{W}^1,p,X)$ and a function of $(\tilde{W}^1,p,X)$. This will almost never be the case because in general the partials of $\tilde{W}^i$ depend on all the variables. Therefore $\delta^1(\tilde{W}^1,p,X)$ and $\delta^2(\tilde{W}^1,p,X)$ must be zero almost everywhere and the solution to (12) in terms of γ^i functions is unique. A more precise statement of why the property implied by (14) is almost never satisfied can be found in the appendix of Blundell, Chiappori, Meghir (2005).

We've showed that we can generically identify the ratios $\frac{\partial F^i}{\partial p}\Big/\frac{\partial F^i}{\partial W^i}$. If we can also identify the ratios $\frac{\partial F^i}{\partial X}\Big/\frac{\partial F^i}{\partial W^i}$, we will have identified the functions F^i up to an increasing transformation. In fact, recovering this second pair of ratios of partial derivatives will be very similar to the first pair.

From the first stage of the household's problem (7) we find that:

$$\frac{\partial V^1(p,r^1,X)}{\partial X}\Big/\frac{\partial V^1(p,r^1,X)}{\partial r^1}+\frac{\partial V^2(p,r^2,X)}{\partial X}\Big/\frac{\partial V^1(p,r^2,X)}{\partial r^2}=\pi$$

Now, using the relationship between W^i and V^i described in equation (6) we find:

$$\frac{\partial W^i(P,y,\tilde{y},X)}{\partial X}\Big/\frac{\partial W^i(P,y,\tilde{y},X)}{\partial \tilde{y}}=\Big(\frac{\partial V^i(p,r^i,X)}{\partial X}\Big)\Big/\Big(\frac{\partial V^i(p,r^i,X)}{\partial r^i}\rho^i_{\tilde{y}}\Big)+\frac{\rho^i_X}{\rho^i_{\tilde{y}}}$$

Then, since $\rho^1=\rho$ and $\rho^2=y-\rho$ we have: $\rho_X=\rho^1_X=-\rho^2_X$ and $\rho_{\tilde{y}}=\rho^1_{\tilde{y}}=-\rho^2_{\tilde{y}}$. Also, recall that earlier we defined the function ψ that gives us the price of the public good π as a function of the exogenous variables and the quantity of the public good X. Therefore we get the following where the right hand side is observable and known:

$$\frac{\partial W_1(P,y,\tilde{y},X)}{\partial X}\Big/\frac{\partial W_1(P,y,\tilde{y},X)}{\partial \tilde{y}}-\frac{\partial W_2(P,y,\tilde{y},X)}{\partial X}\Big/\frac{\partial W_2(P,y,\tilde{y},X)}{\partial \tilde{y}}=\frac{\psi(P,y,\tilde{y},X)}{\rho_{\tilde{y}}} \tag{15}$$

Then, use equation (10) to express the partials of W_i as a function of F^i and $\tilde{W}^i$, plug the expression into (15) and rearrange (this is all exactly the same as for $\frac{\partial F^i}{\partial p}/\frac{\partial F^i}{\partial W^i}$) to get:

$$\frac{\partial F^1}{\partial X}\Big/\frac{\partial F^1}{\partial \tilde{W}^1}\Big(\frac{\partial \tilde{W}^1}{\partial \tilde{y}}\Big)^{-1}-\frac{\partial F^2}{\partial X}\Big/\frac{\partial F^2}{\partial \tilde{W}^2}\Big(\frac{\partial \tilde{W}^2}{\partial \tilde{y}}\Big)^{-1}=\frac{\psi(P,y,\tilde{y},X)}{\rho_{\tilde{y}}}-\frac{\partial \tilde{W}^1}{\partial p}\Big/\frac{\partial \tilde{W}^1}{\partial \tilde{y}}+\frac{\partial \tilde{W}^1}{\partial p}\Big/\frac{\partial \tilde{W}^2}{\partial \tilde{y}} \tag{16}$$

Then, this equation will have a unique solution in terms of the ratios $\frac{\partial F^i}{\partial X}/\frac{\partial F^i}{\partial W^i}$. The argument is exactly the same as above. Therefore, we've generically identified the ratios of partial derivatives of the functions F^i. This implies (see appendix) that we've identified the function F^i up to an increasing transformation. Therefore, we've recovered the collective indirect utility up to an increasing transformation.

Identification of Preferences for the Outside Good

The last remaining object is the preferences for the outside good, $v_i(.)$. Suppose for now that we observe both $h_1(P,y,\tilde{y},X)$ and $h_2(P,y,\tilde{y},X)$ separately. Then, we can recover the indirect utilities $v_i^*(P,y,\tilde{y},X)=v_i(h_i(P,y,\tilde{y},X))$ exactly. More precisely, each choice of $W^i(P,y,\tilde{y},X)$ will determine uniquely the function $v_i^*(P,y,\tilde{y},X)$. Since $W^i(P,y,\tilde{y},X)$ is identified up to an increasing transformation, $v_i^*(P,y,\tilde{y},X)$ is also identified up to the choice of that increasing transformation.

While h_2 only enters the household's problem throught the preferences of agent 2, h_1 enters through the preferences of agent 1 *and* the IC-constraint. Therefore, identifying preferences for the outside good will be different for each agent. However, for both agents identification will rely on the tradeoff between utility from outside consumption $v_i(h_i)$ and within-household consumption $V^i(p, r^i, X)$. While $V^i(p, r^i, X)$ is not identified, note that from (4) we have the following, where $V_2^i(p, r^i, X)$ refers to the partial of $V^i(p, r^i, X)$ with respect to the second argument :

$$V_2^i(p, \rho^i(P, y, \tilde{y}, X), X) = \frac{\partial W^i(P, y, \tilde{y}, X)}{\partial \tilde{y}} \Big/ \rho^i_{\tilde{y}}$$

In the above equation $\rho^i_{\tilde{y}}$ is known. Therefore, each cardinalization of the collective indirect utility $W^i(.)$ will correspond to exactly one $V_2^i(p, \rho^i(P, y, \tilde{y}, X), X)$ function.

The next step is to understand each agent's tradeoff between within-household and outside utility. First, agent 2 will simply equalize the marginal utilities of consumption within the household and outside of the household. The standard condition equating the marginal rate of substitution to the ratio of prices will hold. To see this, simply take the ratio of FOCs with respect to h_2 and with respect to r^2 in the first stage of the household problem defined in (5). We get:

$$\frac{v_2'(h_2)}{V_2^2(p, r^2, X)} = q$$

This condition must hold at the solution of the household's problem. Therefore, we can replace h_2 and r^2 by $h_2(P, y, \tilde{y}, X)$ and $\rho^2(P, y, \tilde{y}, X)$ respectively:

$$v_2'(h_2(P, y, \tilde{y}, X)) = qV_2^2(p, \rho^2(P, y, \tilde{y}, X), X) \tag{17}$$

Given a cardinalization of $W^i(P, y, \tilde{y}, X)$, the right-hand side of the equation above is known. Therefore, we've identified the left-hand side as well. Finding the partial derivatives of $v_2^*(P, y, \tilde{y}, X)$ is now straightforward. Simply multiply the left-hand side of (15) by the relevant partial of $h_2(P, y, \tilde{y}, X)$. That partial is known since we observe $h_2(P, y, \tilde{y}, X)$. For example:

$$\frac{\partial v_2^*(P, y, \tilde{y}, X)}{\partial y} = v_2'(h_2(P, y, \tilde{y}, X)) \frac{\partial h_2(P, y, \tilde{y}, X)}{\partial y}$$

We can recover all the partial derivatives of $v_2^*(P, y, \tilde{y}, X)$ in this way and then simply integrate. Therefore, for a given cardinalization of $W^2(P, y, \tilde{y}, X)$, $v_2^*(P, y, \tilde{y}, X)$ is unique (the integration constant does not affect the household's maximization problem).

For agent 1, the standard condition equating the marginal rate of substitution of within-household consumption and outside consumption to the ratio of prices will not hold.

Instead, for agent 1, the tradeoff between within-household consumption and outside consumption will be determined by the IC-constraint. Intuitively, the IC-constraint tells us that the solution must be such that the increase in agent 1's utility from a marginal increase in $\tilde{y}$ equals the increase in utility if agent 1 were to spend that marginal $\tilde{y}$ only on h_1. If that were not the case, agent 1 would be incentivized to hide part of the income shock and spend it on h_1. More formally, consider the first stage of the household problem once again. The IC-constraint must hold at the solution. That is:

$$W^1(P, y, \tilde{y}, X) + v_1(h_1(P, y, \tilde{y}, X)) = V_0^1 + \frac{1}{q}\int_{y^0}^{\tilde{y}} v_1'((P, y, t, X))\mathrm{d}t$$

Take the partial derivative of the IC-constraint above with respect to $\tilde{y}$ to get:

$$\frac{\partial W^1(P, y, \tilde{y}, X)}{\partial \tilde{y}} + v_1'(h_1(P, y, \tilde{y}, X))(\frac{\partial h_1(P, y, \tilde{y}, X)}{\partial \tilde{y}} - \frac{1}{q}) = 0$$

For a given cardinalization of $W^1(P, y, \tilde{y}, X)$, the only unknown in the equation above is $v_1'(h_1(P, y, \tilde{y}, X))$, which we can therefore solve for. Then, just as for agent 2, we can multiply by the relevant partials derivatives of $h_1(P, y, \tilde{y}, X)$ to find all the partial derivatives of $v_1^*(P, y, \tilde{y}, X) = v_1(h_1(P, y, \tilde{y}, X))$. Finally, we can integrate to find the unique $v_1^*(P, y, \tilde{y}, X)$.

4 Data

4.1 Data Overview

The data used to estimate the model is the Bangladesh Integrated Household Survey. This survey is a large survey of households in Bangladesh. The estimation focuses on households with day laborers because these workers face a great deal of short-term income variation. Therefore, they are particularly likely to be affected by asymmetric information. The sample is restricted to households composed of two adults, the head of the household and the spouse of the head, and their children, grandchildren or nieces and nephews under 25 years old [7]. Only households with at least one child are kept. The purpose of this restriction is to have a homogeneous sample with two main decision makers, which corresponds to the theoretical framework. The restricted sample consists of 1,999 households observed in either 2012 or 2015. In the cases where many household members work as day laborers, the analysis will focus on only one wage earner. Whenever the head of the household works as a day laborer,

[7] All household members that are not the head or the spouse of the head are referred to as children from now on.

the analysis will focus on him. In the 4% of cases where the spouse works as a day laborer and the head does not the analysis focuses on her.

Table 1: Day Laborer Summary Statistics

	Mean	SD
Male	0.96	0.20
Age	41.5	10.6
Number of Children	2.25	1.14
No Education	0.64	0.48
Rural	0.94	0.24
Weekly Income ($ US)	12.9	6.81
Weekly Household Income($ US)	18.3	11.6
Days Worked in the Past Week	4.68	1.86
Weekly Hours	36.2	16.8
Works in Agriculture	0.72	0.45
Works In Own Village	0.89	0.31
Observations	1999	

Table 1 presents some summary statistics for the day laborers and the households that compose the main estimation sample. The day laborers are overwhelmingly male. They are part of very poor rural households: the average weekly wage is less than 13 $US, and more than half of them have no education whatsoever. A large majority work as agricultural day laborers. Other typical industries include construction and transport. The typical arrangement in these households (Khandker and Mahmud 2012) is that the head of the household works as a day laborer, while the spouse is in charge of generating some additional income through home agriculture or raising poultry. While the male wage earner that works outside of the household earns more on average, his income tends to be variable and depend heavily on the seasonal availability of work.

4.2 Crucial Data Features and Variable Description

Two non-standard features of the BIHS are going to be crucial to estimate the model. First, for day laborers, the survey reports both average weekly income and income in the seven days preceding the survey. Having these two measures of income is necessary to construct a measure of short-term income shocks, which is the income variation that is potentially affected by asymmetric information. Second, food consumption is reported at the individual level. Therefore, it will serve as the assignable good that is necessary to identify the model. Importantly, individual food consumption is reported in the 24 hours preceding the survey.

The model describes how short-term income shocks affect short-term consumption shares in the presence of asymmetric information. Therefore, it is important to use a short-term, non-durable measure of consumption – such as food consumption in the past 24 hours – as the assignable good. Clothing, which has been commonly used as an assignable good in the literature, would not be adequate here.

Three categories of variables are going to be necessary to estimate the model: consumption variables, income variables, and price variables. Starting with consumption variables, the adult household members will consume two private goods: food c_i and a non-food composite private good x_i. The BIHS contains a food consumption module that is going to allow the construction of an individual level food consumption variable c_i. The module breaks down individual food consumption at each meal taken in the household in the 24 hours preceding the survey. Only meals in which the day laborer participates are kept to avoid issues with eating outside of the household. This module can be combined with data on food purchased by the household in the last week to get a precise measure of the food expenditures for each household member. The non-food private composite good x_i is constructed by summing expenditures on transport, communications, energy, hygiene, and cosmetics. This composite good is observed at the household level: $x_1 + x_2$ is observed but not x_1 and x_2 separately. The adult household members also derive utility from a public good X: expenditures on children. This variable is constructed as the sum of expenditures on food consumed by children, clothes for children and education expenditures.

As noted above, the survey asks day laborers to report both their average weekly income (denoted $\bar{y}$) and their income over the past seven days (denoted y_s). From these two measures of income, the income shock $\tilde{y}$ is constructed as the ratio of the two: $\tilde{y} = \frac{y_s}{\bar{y}}$. Total household income Y is constructed as the sum of the day laborer's average weekly income and other income sources including wage income from other household members, income from selling agricultural products and remittances. Figure 1 shows histograms of the day laborer's two income measures, as well as $\ln(\tilde{y})$, which is the way in which the income shock will enter the empirical specification. The income shock $\ln(\tilde{y})$ is quite symmetrically distributed around 0. Note that the distribution of income in the last seven days is slightly to the right of the distribution of average weekly income. This is explained by the survey only recording information on day laborers if they had positive income in the past seven days. Most day laborers work in agriculture and are affected by seasonal availability of jobs. Since they are not necessarily working all year long, conditional on having worked in the past seven days, their income in the past seven days is higher than their average weekly income.

There is some sample selection based on whether the day laborer worked or not in

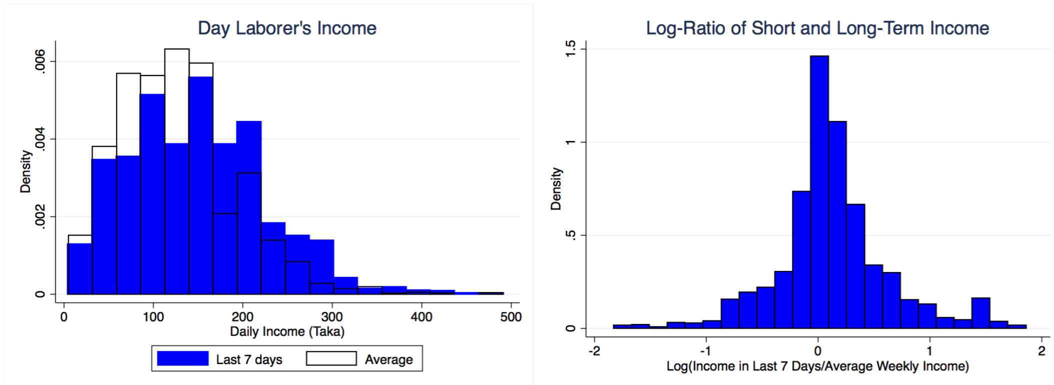

Figure 1: Day Laborer's Income

the past seven days. However, conditional on the probability of working in a given seven day period, whether or not a day laborer worked in the seven days before the survey is random. Similarly, conditional on the probability of working in a given seven day period, the income shock depends on the date of the survey interview. There is no reason to think that interviewed day laborers would manipulate their income in the week before the interview or that the interviewers would systematically choose the date of the interview in any particular way. Therefore, conditional on the probability of working in a given seven day period, the income shock can be thought of as random. In an effort to control for the unobserved probability of being in the sample, one of the empirical specifications will include an indicator variable for agricultural day laborers. The idea is that agricultural day laborers have more seasonal fluctuation in job availability and therefore are more likely to be affected by sample selection. The results, which are shown later, are unaffected by the inclusion of this indicator variable, which indicates that sample selection is not driving the results.

The final variables required for estimation are prices. The price of the composite private good x_i is normalized to 1. A Laspeyres price index for food is constructed at the village level, using food items that are consumed by at least 10% of households. The weight on the price of each food item is the average consumed quantity of that item across the whole sample. Similarly, a clothing price index is constructed at the village level. The price of children consumption is then a weighted average of the food price index and the clothing price index. Once again, the weights are the average share of children expenditures on food and clothing respectively. Food expenditures and children expenditures are then divided by their respective price indexes to get the final consumption measures. Table 2 presents summary statistics for the variables discussed in this section.

Table 2: Consumption, Income and Prices

	Mean	SD
Consumption		
Day Laborer Food (c_1)	39.7	17.0
Home Producer Food (c_2)	32.7	15.4
Non-food Private Good ($x_1 + x_2$)	37.3	16.4
Children Consumption (X)	56.1	46.8
Total Expenditures (E)	169.7	73.9
Income		
Income in Last 7 Days ($\tilde{y}$)	147.3	77.8
Average Weekly Income ($\bar{y}$)	128.8	66.4
Total Household Income (Y)	209.1	132.7
Prices		
Food Price (p)	1.00	0.052
Children Consumption Price (π)	1.00	0.045
Observations	1999	

All values in Taka/day. 75 Taka ≈ 1 USD.

5 Empirical Specification

Although the model is identified non-parametrically, precise estimates require reducing the dimensionality through a parametric specification. The parametric specification is similar to Cherchye, De Rock, Vermeulen's (2012). The indirect utility from the consumption of food c_i and the non-food composite private good x_i, conditional on expenditures on children X is assumed to be:

$$V^i(p, \rho^i, X) = \frac{\ln(\rho^i) - \alpha^i \ln(p)}{p^{\beta^i}} + \kappa^i \ln(X),$$

where $\alpha^i = \alpha^i_0 + \alpha^i_1 K$ and $\kappa^i = \kappa^i_0 + \kappa^i_1 K$ with K the number of children in the household. In principle, α^i, β^i and κ^i could be functions of demographics. However, in order to avoid overspecification, only number of children is allowed to impact α^i and κ^i. As a reminder, p is the price of food consumption c_i and ρ^i is the allocated budget to agent i to spend on private goods. The preferences for the private good correspond to the preferences underlying Deaton and Muellbauer's (1980) Almost Ideal Demand System[8]. Using Roy's identity, the

[8]The notation α^i and β^i in this paper is consistent with the notation in Deaton and Muellbauer's (1980) seminal paper.

Marshallian demand for food consumption c_i can be derived.

$$c_i = [\alpha^i + \beta^i(\ln(\rho^i) - \alpha^i \ln(p))]\frac{\rho^i}{p} \tag{18}$$

Then, using the budget constraint from each agent's second stage the demand for non-food consumption x_i can be derived.

$$x_i = \rho^i - pc_i = [(1 - \alpha^i) - \beta^i(\ln(\rho^i) - \alpha^i \ln(p))]\rho^i \tag{19}$$

The demand for the public good can be derived from the first stage of the household's problem:

$$X = \kappa^1 \frac{\rho^1 p^{\beta^1}}{\pi} + \kappa^2 \frac{\rho^2 p^{\beta^2}}{\pi} \tag{20}$$

These demand functions demand depend on the sharing rule ρ^i. The sharing rule is determined in the first stage by the efficiency rule:

$$(\mu + \lambda_2)\frac{\partial V^1(p, \rho^1, X)}{\partial \rho^1} = (1 - \mu)\frac{\partial V^2(p, \rho^2, X)}{\partial \rho^2}$$

Using the fact that $\frac{\partial V^1(p,\rho^i,X)}{\partial \rho^i} = \frac{1}{\rho^i p^{\beta^i}}$ and that $\rho = \rho^1 = y - \rho_2$ we get the following expression for the sharing rule ρ:

$$\rho = \frac{y}{1 + p^{(\beta_1 - \beta_2)} \frac{(1-\mu)}{(\mu + \lambda_2)}} \tag{21}$$

Following Browning, Chiappori and Lewbel (2008) the decision power of agent 1 for a given realization $\tilde{y}$ is assumed to take a logistic form:

$$\frac{\mu + \lambda_2}{1 + \lambda_2} = \frac{\exp(\gamma_1 \ln(\tilde{y}) + \gamma_2 \ln(\bar{y}) + \gamma_3 \ln(E) + \gamma_4 \ln(Y) + \gamma_D D)}{1 + \exp(\gamma_1 \ln(\tilde{y}) + \gamma_2 \ln(\bar{y}) + \gamma_3 \ln(E) + \gamma_4 \ln(Y) + \gamma_D D)} \tag{22}$$

Note that under perfect information we would expect the decision power not to vary with income shocks $\tilde{y}$ and therefore we would have $\lambda_2 = 0$ and $\gamma_1 = 0$. The decision power is allowed to depend on the wage earner's average income $\tilde{y}$. In general, the decision power can also vary across households. In order to capture some of that variation, the decision power can depend on a vector of demographic variables D and on total household income Y.

The decision power is also allowed to depend on total household expenditures E: the sum of all private and public expenditures in the household. This variable is included to capture potentially different risk aversions between the two adults in the household. For example, if the wage earner is less risk averse than the home producer, insurance within the household would imply that the wage earner's consumption share be greater when total household expenditures E are high. short-term total household expenditures E and short-term income shock $\tilde{y}$ of the wage earner are likely to be correlated. Therefore, controlling for E is important to avoid a bias in $\hat{\gamma}_1$ and potentially overestimating the effect of asymmetric information.

The parametric specification of the decision power in equation (22) gives:

$$\frac{\mu + \lambda_2}{1 - \mu} = \exp(\gamma_1 \ln(\tilde{y}) + \gamma_2 \ln(\bar{y}) + \gamma_3 \ln(E) + \gamma_4 \ln(Y) + \gamma_D D)$$

Plugging this into equation (21) gives us an expression for ρ as a function of variables in the data:

$$\rho = y \times \frac{\exp((\beta_2 - \beta_1)\ln(p) + \gamma_1 \ln(\tilde{y}) + \gamma_2 \ln(\bar{y}) + \gamma_3 \ln(E) + \gamma_4 \ln(Y) + \gamma_D D)}{1 + \exp((\beta_2 - \beta_1)\ln(p) + \gamma_1 \ln(\tilde{y}) + \gamma_2 \ln(\bar{y}) + \gamma_3 \ln(E) + \gamma_4 \ln(Y) + \gamma_D D)} \tag{23}$$

Finally, this equation for ρ can be plugged into the demand equations (18), (19), and (20) to get closed form solutions for demands as functions of data and parameters.

This system of equations can then be estimated. To be more precise, we will model $(c_1, c_2, x = x_1 + x_2, X)$ as observable functions of $(p, \pi, y, \tilde{y}, \bar{y}, E, Y, D)$. Additive errors are added to the demand equations to account for unobservable heterogeneity. The system of equations is estimated with a feasible generalized least squares estimator. Note that this estimator allows for correlated errors across the different goods.

6 Results and Counterfactuals

6.1 Main Result: the Effect of Asymmetric Information

Since perfect information is a special case of the model, the estimation results provide a natural test for asymmetric information. Under the null hypothesis of perfect information, the effect of the income shock should be zero. Therefore, if the effect of an income shock on the decision power is positive and statistically significant, we can reject perfect information.

The main estimation results are given in Table 3. The first and second columns correspond to the estimated coefficients and corresponding standard errors when only total household income Y is included in the decision power. The third and fourth columns correspond to the case where two other demographic variables are included: an indicator variable for whether the day laborer works in agriculture and the number of kids in the household.

Table 3: Structural Estimation Results

	Baseline specification		Demographic Controls	
	Coef.	SE	Coef.	SE
Decision Power				
γ_1[Income Shock]	0.030*	0.0099	0.031*	0.0098
γ_2[Average Income]	0.031*	0.0089	0.031*	0.0089
γ_3[Tot. Expenditures]	-0.20*	0.014	-0.21*	0.014
γ_4[Total HH Income]	0.34*	0.015	0.32*	0.016
γ_{D1}[Works in Agr.]			0.011	0.0095
γ_{D2}[Number of Kids]			0.12*	0.022
Private Preferences				
α_0^1[Constant]	0.84*	0.032	0.82*	0.032
α_1^1[Number of Kids]	-0.011*	0.0017	-0.023*	0.0028
β^1	-0.060*	0.0046	-0.052*	0.0044
α_0^2[Constant]	0.15*	0.068	-0.035	0.093
α_1^2[Number of Kids]	-0.053*	0.0049	0.059*	0.025
β^2	0.26*	0.037	0.27*	0.039
Public Preferences				
κ_0^1[Constant]	0.0061	0.078	0.059	0.065
κ_1^1[Number of Kids]	0.18*	0.030	0.15*	0.022
κ_0^2[Constant]	-0.068	0.19	-0.35*	0.17
κ_1^2[Number of Kids]	0.36*	0.070	0.54*	0.065
N	1999		1999	
1 SD Income Shock	0.49 p.p.		0.48 p.p.	

FGNLS estimator. * Significant at 5% level.

The first thing to note is that the estimated coefficient of the income shock $\tilde{y}$ is, in fact, positive and statistically significant. This allows us to reject perfect information and provides evidence that these households are affected by asymmetric information. The last row of the table quantifies the effect of an income shock: in the baseline specification, a one standard deviation income shock above the mean increases the consumption share of the wage earner by 0.49 percentage points.

The baseline specification and the specification with the additional demographic controls are extremely similar for most coefficients. In particular, the addition of these demographic controls does not affect how the decision power varies with the income shock. The fact that the estimation is not sensitive to the inclusion of demographic controls provides some evidence that the results are not being driven by unobserved heterogeneity across households.

The estimation of the preference parameters suggest that wage earners and home producers have different preferences. The wage earner prioritize food consumption, as can be seen by the large α_0^1 relative to α_0^2. The home producer also has a stronger preference for expenditures on children, consistent with previous results in the literature.

6.2 Work Location Heterogeneity

Some households are likely to be more affected by asymmetric information than others. In particular, if the wage earner works further from home, income shocks are likely to be less observable to other household members. Then, in order to be incentivized to report truthfully, such a wage earner's consumption share would be more responsive to income shocks.

Analyzing heterogeneity in the response of consumption shares to income shocks serves two main purposes. First, it provides a test for whether the asymmetric information mechanism is driving the main results. The prediction under asymmetric information is that the response of consumption shares to income shocks would be larger for wage earners working far from home relative to those living close to home. The heterogeneity analysis tests this prediction. Second, identifying households that are likely to be more affected by asymmetric information could help target public policies aimed at reducing the effect of asymmetric information.

The work location heterogeneity is modeled by allowing the effect of income shocks on the decision power to be different for two groups: those that work far from their home and those that work near their home. Specifically, 11% of wage earners work outside of the union in which they live[9]. The model is re-estimated allowing the effect of income shocks on decision power to be different for these wage earners.

More formally, d is defined to be equal to 1 if the wage earner works outside of the union in which they live, 0 otherwise. The main empirical specification is augmented by having $\gamma_1 = \gamma_{10} + \gamma_{11} \times d$ and including d additively in the decision power equation. This is the only change to the model in the baseline heterogeneity analysis. In a second analysis, in addition to the baseline changes, the effect of other variables on decision power is also allowed to differ for the two subsamples by having $\gamma_i = \gamma_{i0} + \gamma_{i1} \times d$ for $i = 2, 3, 4$.

The results are presented in Table 4. The first two column correspond to the case where only the effect of the income shock is allowed to vary across the two subsamples. The second pair of columns correspond to the case where all the variables are interacted.

[9] Unions are the smallest administrative unit in Bangladesh. There are 4562 unions in Bangladesh, with an average area of 32 square kilometers.

Table 4: Work Near Home vs. Work Far From Home

	Shock Interaction		All Interactions	
	Coef.	SE	Coef.	SE
Work Near ($d = 0$)				
γ_{10}[Income Shock]	0.018	0.0100	0.022*	0.010
γ_{20}[Average Income]	-0.0038	0.011	-0.013	0.012
γ_{30}[Tot. Expenditures]	-0.20*	0.013	-0.19*	0.014
γ_{40}[Total HH Income]	0.35*	0.015	0.36*	0.015
Work Far ($d = 1$)				
γ_{11}[Income Shock]	0.10*	0.022	0.081*	0.033
γ_{21}[Average Income]			0.045	0.031
γ_{31}[Tot. Expenditures]			-0.055	0.038
γ_{41}[Total HH Income]			-0.048	0.033
γ_{01}[Constant]	-0.45*	0.11	-0.024	0.22
N	1999		1999	

FGNLS estimator. * Significant at 5% level.

The results provide strong evidence in favor of the asymmetric information mechanism presented in this paper. Consistent with the effect of asymmetric information the wage earner's consumption increases significantly more with income shocks for those that work far from their home. In fact, the magnitude of the effect is 4 to 5 times larger depending on the specification. In addition, note that the effect of the other variables on decision is not statistically different for the the two subsamples of wage earners. This helps address the concern that the difference in the effect of income shocks is being driven by selection or preference heterogeneity.

The large difference in the effect of income shocks for the two subsamples suggests that households are affected quite heterogeneously by asymmetric information. As suggested by the results in table 2, this heterogeneity is likely to be the result of variation in how observable income shocks are to other household members. In addition, variation in altruism or preferences for hiding income might also be important dimensions of heterogeneity to consider. Overall, this suggests the importance of targeting policies to reduce asymmetric information appropriately.

6.3 Daily Wage and Days Worked Variation

A possible alternative explanation for the previous empirical results is that consumption shares and income shocks are correlated because more resources are needed by the wage earner when working more. This is particularly a concern because the assignable good being used to identify the decision power is food consumption. Since most of the wage earners work in manual labor industries, it is likely that when they work more they have higher nutritional needs.

In order to show that nutritional needs are not the main explanation for the empirical results, income variation is separated between daily wage and days worked variation. The prediction is that, if nutritional needs are driving the results, the variation in consumption shares should be explained by the variation in days worked. If, on the contrary, asymmetric information is driving the results, the variation in consumption shares should be explained by the variation in daily wage. The reasoning is that nutritional needs of a worker will depend more on the number of days worked than on the daily wage. In addition, daily wage variation is likely to be less observable by other household members than how many days the wage earner worked. Comparing the effect of wage variation and days worked variation could therefore serve as a test of the asymmetric information mechanism, even in the absence of nutritional needs. Here, the comparison serves the added purpose of ruling out an alternative mechanism.

Separating daily wage and days worked variation requires observing measures of these variations. The measure of the short-term income shock that has been used up to now is the ratio of income over the last seven days (y_s) and average weekly income ($\bar{y}$): $\tilde{y} = \frac{y_s}{\bar{y}}$. Income over the last seven days was constructed from the survey data as the product of number of days worked in the last seven days (y_d) and average daily wage in the last seven days (y_w): $y_s = y_d \times y_w$. Since the income shock enters the decision power expression through a logarithmic function it is natural to divide the wage and days variation using:

$$\ln(\tilde{y}) = \ln(\frac{y_w \times y_d}{\bar{y}}) = \ln(y_w) + \ln(y_d) - \ln(\bar{y})$$

Then we can re-estimate the model by including $\gamma_{1w} \ln(y_w) + \gamma_{1d} \ln(y_d)$ in the expression for the decision power instead of just $\gamma_1 \ln(\tilde{y})$ [10]. If the results are not being driven by nutritional needs, we would expect a large and positive γ_{1w}.

Figure 2 plots the histograms of the two sources of income variation. The main takeaway is that days worked varies quite a bit and does not seem to follow the usual 5 day work week that is common in developed countries. This is an important point because without significant days worked variation it would be impossible to compare the effect of the two sources of income variation.

The estimation results for the decision power are given in Table 5. They rule out nutritional needs as the driver of the empirical results in the main specification. Consistent with the

[10] Since $\gamma_2 \ln(\bar{y})$ is already in the main specification, there is no need to include another term for it. Simply note that in the new estimation the coefficient on $\ln(\bar{y})$ will correspond to $\gamma_2 - \gamma_1$ from the original specification.

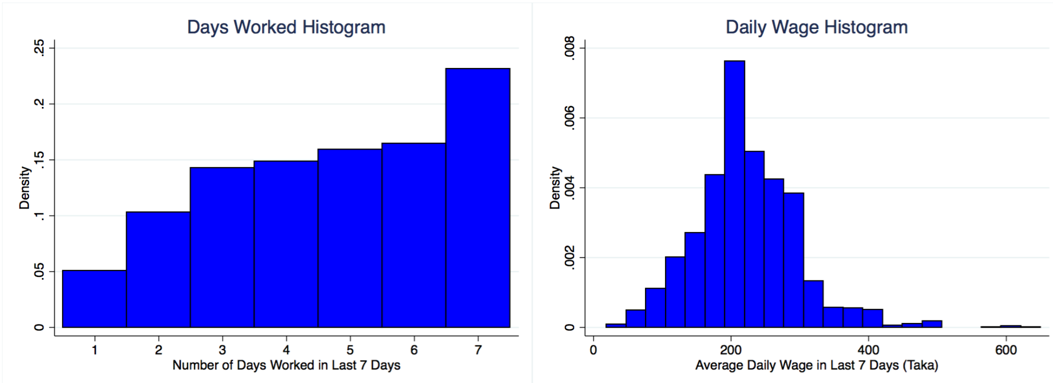

Figure 2: Daily Wage and Days Worked Variation

asymmetric information mechanism, the coefficient on daily wage is positive and significant. In fact, the estimated coefficient for daily wage is twice as large the one for days worked. The last two rows allow a more intuitive comparison of the magnitudes. A daily wage one standard deviation above the mean leads to a 0.33 percentage point increase in the wage earner's consumption share. Similarly, a one standard deviation increase in the days worked leads to a 0.17 percentage point increase in the wage earner's consumption share. This clearly shows that the variation in consumption share in response to income shocks results from daily wage variation, more so than from days worked variation. Therefore, the analysis is consistent with asymmetric information as the main mechanism, rather than nutritional needs.

Table 5: Daily Wage vs. Days Worked Variation

	.	
	Coef.	SE
γ_{1w}[Daily Wage]	0.050*	0.015
γ_{1d}[Days Worked]	0.022*	0.011
γ_2[Average Income]	-0.00029	0.011
γ_3[Tot. Expenditures]	-0.20*	0.014
γ_4[Total HH Income]	0.34*	0.016
N	1999	
$\gamma_{1w} - \gamma_{1d}$ (pval)	0.28	0.069
1 SD Shock (Wage)	0.33 p.p.	
1 SD Shock (Days)	0.17 p.p.	

FGNLS estimator. * Significant at 5% level.

6.4 Counterfactual: A Naive Approach to Asymmetric Information

The model presented in this paper emphasizes that cooperative households can reduce the cost of asymmetric information by incentivizing wage earners to reveal their true income. The empirical results provide evidence that the households do in fact behave in this way. However, this raises the question: what would be the welfare cost for households if they did not incentivize truthful income reporting? This question is important because it can help shed light on the experimental results on asymmetric information. Experiments are typically one-offs and situations household members have no experience dealing with. As a result, household members are likely not able to coordinate to provide incentives in an experiment, as they would in a repeated setting outside of the lab. Therefore, experiments might overstate the cost of asymmetric information.

Wage earners that are not incentivized will choose to hide some income. How much will depend on their preferences for hiding income. Therefore, quantifying the welfare cost for a "naive" household of not incentivizing the wage earner requires estimating preferences for hiding income. The identification result states that these preferences can be estimated if a hideable good is observed at the individual level. A hideable good is simply a good that can be consumed without other household members' knowledge. Here the hideable good will be betel nuts, a widely used stimulant in Bangladesh, similar to tobacco. Individual level consumption of betel nuts is observed. 671 out of the 1999 day laborers consume betels nuts and therefore the counterfactual analysis will focus on this subsample.

The counterfactual simulation then proceeds in two steps. In the first step, preferences for betel nuts (denoted h_1) are estimated. The parametric specification is chosen to be $v^1(h_1) = \tau_1 \ln(h_1)$. A closed form solution for h_1 as a function of parameters can be derived by combining the first order condition for the hideable good (24) and the partial derivative of the incentive compatibility constraint from the first stage of the household's problem with respect to $\tilde{y}$ (25):

$$v_1'(h_1) = \frac{\partial V^1(p, \rho^1, X)}{\partial \rho^1} + \frac{\lambda_2}{\mu + \lambda_2} \frac{1}{\frac{\partial h_1}{\partial \tilde{y}}} v_1'(h_1) \tag{24}$$

$$v_1'(h_1) = \frac{\partial V^1(p, \rho^1, X)}{\partial \rho^1} \frac{\partial \rho^1}{\partial \tilde{y}} + \frac{\partial h_1}{\partial \tilde{y}} v_1'(h_1) \tag{25}$$

The derivations are tedious and not particularly insightful and therefore given in the appendix. We get:

$$h_1 = \tau_1 \frac{(\gamma_1 + \tilde{y}) - \sqrt{(\gamma_1 + \tilde{y})^2 - 4\frac{\mu}{\mu+\lambda_2}\gamma_1\tilde{y}}}{2\gamma_1\tilde{y}}$$

Estimating τ_1 using this equation would require knowing γ_1 and the other parameters that go into μ and λ_2. Obviously, the true values are unknown. However, since these parameters have already been consistently estimated in the main estimation results, the estimated values can be plugged in instead. τ_1 can then be consistently estimated using a linear least squares estimator. However, the two-stage estimation means that the standard errors for τ_1 have to be bootstrapped[11].

The second step of the counterfactual is to simulate the behavior of wage earners that are not incentivized to reveal their income. Without incentives, the wage earner will decide how much to hide based on the preferences for hiding income relative to revealing income. This choice is determined by the following first-order condition:

$$v_1'(h_1) = \frac{\partial V_1(p, \rho^1, X)}{\partial \rho^1}\frac{\partial \rho^1}{\partial y}$$

The intuition is that the wage earner will choose to hide until the marginal utility from hiding is equal to the marginal utility from reporting truthfully. The marginal utility from reporting truthfully is equal to the product of the marginal increase in the wage earner's share of resources and the marginal utility increase from the extra resources. Given the functional forms, this gives us:

$$h_1^c = \frac{p_1^\beta y}{\tau_1}$$

where h_1^c is the counterfactual value of the betel nut consumption. This equation determines the maximum the wage earner would like to hide. In order to have a more realistic simulation, the wage earner is constrained to hiding a maximum of 10% of income over the past seven days in the baseline simulation and 20% in a second simulation. This reflects the fact that only part of the income is unobservable to other household members and can be hidden.

The results are presented in Table 6. When wage earners can hide up to 10% of income, they hide on average 44% of that 10%. 15% choose to hide the maximum amount they can hide. When they can hide up to 20% of income, they hide on average 26% of that 20% and 4.3% choose to hide the maximum amount they can hide.

Interestingly, wage earners choose to report a large portion of the income they could

[11] I have not finalized the bootstrap. For now I report standard errors from the linear estimation, ignoring the two-stage nature of the estimation.

Table 6: Counterfactual Analysis Results

Unobservable Income	τ_1 (SE)	Proportion Hidden	Share hiding 100%
10%	0.11* (0.004)	44%	15.1%
20%	0.11* (0.004)	26%	4.3%

have hidden, even in the absence of incentives. This reflects a preference for goods that are consumed in the household, such as food and other essentials. It is also consistent with the relative small consumption share changes that where estimated in the main specification: a 0.49 percentage point increase in wage earners' consumption share is enough for them to report an income shock one standard deviation above the mean.

The welfare implications of not incentivizing truthful income reporting are concentrated on the spouse and children. By the nature of the incentive compatibility constraint, the wage earner is indifferent between being incentivized or not. However, the counterfactual shows that when 10% of income can be hidden, incentives increase the welfare of the spouse and children equivalently to a 4.4% increase in total household income. When the wage earner can hide up to 20% of income, that number goes up to 5.2%. Therefore, the welfare gains from providing incentives are considerable.

7 Conclusion

This paper shows that asymmetric information about income matters for household decisions, resource-sharing and welfare. A new model of a household facing asymmetric information about the income of the wage earner provides a theoretical framework to predict the effect of asymmetric information on household outcomes and to conduct counterfactual analyses. Then, a new identification result, specific to this model, shows that individual welfare functions can be estimated from consumption data. Finally, the model is estimated using a sample of Bangladeshi day laborers. The estimation provides evidence that the day laborers are affected by asymmetric information but that cooperative households are able to significantly reduce the cost of asymmetric information.

References

Ambler, K. (2015). Don't tell on me: Experimental evidence of asymmetric information in transnational households. *Journal of Development Economics*, *113*, 52–69. Retrieved

March 3, 2020, from https://econpapers.repec.org/article/eeedeveco/v_3a113_3ay_3a2015_3ai_3ac_3ap_3a52-69.htm

Ashraf, N. (2009). Spousal control and intra-household decision making: An experimental study in the philippines. *American Economic Review*, *99*(4), 1245–1277. https://doi.org/10.1257/aer.99.4.1245

Blundell, R., Chiappori, P.-A., & Meghir, C. (2005). Collective labor supply with children. *Journal of political Economy*, *113*(6), 1277–1306.

Boltz, M., Marazyan, K., & Villar, P. (2019). Income hiding and informal redistribution: A lab-in-the-field experiment in senegal. *Journal of Development Economics*, *137*, 78–92. https://doi.org/10.1016/j.jdeveco.2018.11.004

Brown, C., Calvi, R., & Penglase, J. (2018). Sharing the pie: Undernutrition, intra-household allocation, and poverty. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3199129

Castilla, C., & Walker, T. (2013). Is ignorance bliss? the effect of asymmetric information between spouses on intra-household allocations. *American Economic Review*, *103*(3), 263–268. https://doi.org/10.1257/aer.103.3.263

Chiappori, P.-A., & Ekeland, I. (2009). The microeconomics of efficient group behavior: Identification 1. *Econometrica*, *77*(3), 763–799.

Chiappori, P.-A. (1988). Rational household labor supply. *Econometrica*, *56*(1), 63–90. https://doi.org/10.2307/1911842

Chiappori, P.-A. (1992). Collective labor supply and welfare. *Journal of Political Economy*, *100*(3), 437–467. Retrieved September 24, 2019, from https://www.jstor.org/stable/2138727

Dagnelie, O., & Lemay-Boucher, P. (2012). Rosca participation in benin: A commitment issue*: Rosca participation in benin. *Oxford Bulletin of Economics and Statistics*, *74*(2), 235–252. https://doi.org/10.1111/j.1468-0084.2011.00641.x

Dercon, S., & Krishnan, P. (2000). In sickness and in health: Risk sharing within households in rural ethiopia. *Journal of Political Economy*, *108*(4), 688–727. https://doi.org/10.1086/316098

Dubois, P., & Ligon, E. (2011). *Incentives and nutrition for rotten kids: Intrahousehold food allocation in the philippines*.

Green, J. R., & Laffont, J.-J. (1986). Partially verifiable information and mechanism design. *The Review of Economic Studies*, *53*(3), 447. https://doi.org/10.2307/2297639

Hall, R. E., & Mishkin, F. S. (1982). The sensitivity of consumption to transitory income: Estimates from panel data on households. *Econometrica*, *50*(2), 461–481. https://doi.org/10.2307/1912638

Hoel, J. B. (2015). Heterogeneous households a within-subject test of asymmetric information between spouses in kenya. *Journal of economic behavior & organization*, *118*, 123–135. https://doi.org/10.1016/j.jebo.2015.02.016

Khandker, S. R., & Mahmud, W. (2012). *Seasonal hunger and public policies: Evidence from northwest bangladesh*. The World Bank. https://doi.org/10.1596/978-0-8213-9553-0

Klein, M. J., & Barham, B. L. (2018). *Point estimates of household bargaining power using outside options*. University of Wisconsin, Agricultural and Applied Economics.

Ligon, E. A. (2011). Dynamic bargaining in households (with an application to bangladesh). *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1776810

Malapit, H. J. L. (2012). Why do spouses hide income? *The Journal of Socio-Economics*, *41*(5), 584–593. https://doi.org/10.1016/j.socec.2012.04.007

Mazzocco, M. (2007). Household intertemporal behaviour: A collective characterization and a test of commitment. *The Review of Economic Studies*, *74*(3), 857–895.

Munshi, K., & Rosenzweig, M. (2016). Networks and misallocation: Insurance, migration, and the rural-urban wage gap. *American Economic Review*, *106*(1), 46–98. https://doi.org/10.1257/aer.20131365

Squires, M. (2020). Kinship taxation as an impediment to growth: Experimental evidence from kenyan microenterprises, 74.

Appendix

We will show that the knowledge of the ratio of partial derivatives of a function allows us to recover that function up to an increasing transformation. We will show this for a function with three arguments because this is the relevant case in this paper. However, this argument can be generalized to an arbitrary function.

We are interested in the function $W(x, y, z)$. Suppose that the function is monotonic in each one of it's arguments. If it is not we can apply the argument below over monotonic regions of the function. Denote W_ξ for $\xi = x, y, z$ the partial derivatives of this function. We know $r^1(x, y, z) = W_y/W_x$ and $r^2(x, y, z) = W_z/W_x$ the ratios of partial derivatives. The question is: if two functions $W(x, y, z)$ and $\hat{W}(x, y, z)$ have the same ratios of partial derivatives, what is the relationship between of $W(x, y, z)$ and $\hat{W}(x, y, z)$? The goal is to show that it must be that $\hat{W}(x, y, z) = G(W(x, y, z))$ for some increasing function $G(.)$.

First note that $W_y/W_x = \hat{W}_y/\hat{W}_x$ and $W_z/W_x = \hat{W}_z/\hat{W}_x$ imply $\hat{W}_x/W_x = \hat{W}_y/W_y = \hat{W}_z/W_z \equiv g(x, y, z)$ for some function $g(x, y, z)$. Now we have that $\hat{W}_\xi = g(x, y, z)W_\xi$ for $\xi = x, y, z$.

The next step is to integrate the partial derivatives of $\hat{W}(x,y,z)$:

$$\hat{W}(x,y,z) = \int \hat{W}_x(x,y,z)dx = \int g(x,y,z)W_x(x,y,z)dx$$

Then apply the implicit function theorem to the relationship $W(x,y,z) = W$ to define the function $h(y,z,W)$ such that $h(y,z,W(x,y,z)) = x$. Then we can write the equation above as:

$$\hat{W}(x,y,z) = \int g(h(y,z,W(x,y,z)),y,z)W_x(x,y,z)dx$$

Then we can integrate by substitution to find:

$$\hat{W}(x,y,z) = \int g(h(y,z,W),y,z)dW = H^1(y,z,W(x,y,z)) + k^1(y,z) \tag{A1}$$

Simlarly, we find:

$$\hat{W}(x,y,z) = H^2(x,z,W(x,y,z)) + k^2(x,z) \tag{A2}$$

$$\hat{W}(x,y,z) = H^3(x,y,W(x,y,z)) + k^3(x,y) \tag{A3}$$

Then we can take the partial derivatives of (A1), (A2), (A3) with respect to x, y, z to find the following relationships:

$$\begin{aligned}
\hat{W}_x &= g(x,y,z)W_x = H_3^1 W_x = H_3^2 W_x + H_1^2 + k_1^2 = H_3^3 W_x + H_1^3 + k_1^3 \\
\hat{W}_y &= g(x,y,z)W_y = H_3^2 W_y = H_3^1 W_y + H_1^1 + k_1^1 = H_3^3 W_y + H_2^3 + k_2^3 \\
\hat{W}_z &= g(x,y,z)W_z = H_3^3 W_z = H_3^1 W_z + H_2^1 + k_2^1 = H_3^2 W_z + H_2^2 + k_2^2
\end{aligned}$$

The first thing to notice is that $H_3^1 = H_3^2 = H_3^3 = g(x,y,z)$. This, in turn, implies that $H_i^j = -k_i^j$ for $i = 1,2$ and $j = 1,2,3$. Notice that the k functions only depend on two arguments, while the H functions depend on all three. Therefore, we find that the partial derivative of the H functions with respect to the first two arguments do not depend on the variable that only enters the H functions through W. For example $H_1^1(x,y,z) = -k_1^1(y,z)$ and $H_2^1(x,y,z) = -k_2^1(y,z)$. Therefore, the partial derivatives of H^1 with respect to the first two arguments do not depend on x. This, in turn implies the function H^1 is additively separable between the first two and the the third argument. In other words, the function takes the form:

$$H^1(y,z,W(x,y,z)) = \tilde{H}(W(x,y,z)) + l^1(y,z)$$

In addition, we also get that $H_1^1(y,z,W(x,y,z)) = l_1^1(y,z) = -k_1^1(y,z)$ and $H_2^1(y,z,W(x,y,z)) = l_2^1(y,z) = -k_2^1(y,z)$. Therefore, the sum $l^1(y,z) + k^1(y,z)$ is a constant that does not depend on y or z. Putting this into equation (10) we find that $\hat{W}(x,y,z) = \tilde{H}(W(x,y,z)) + k =$

$G(W(x, y, z))$ for some function $G(.)$. Therefore, we have shown that the ratios of partial derivatives allow us to recover the function up to some transformation.

In addition, if we know the sign of one of the partial derivatives of W, say for example $W_x > 0$, then this implies that $G(.)$ has to be an increasing function. This would be the case in this paper because it makes sense to assume that preferences are such that collective utilities are decreasing in p for example.