Sharing wage risk

S. Bonhomme, P.A. Chiappori, R. Townsend, H. Yamada*
CEMFI
Columbia University
Massachusetts Institute of Technology
Osaka University

1 Introduction

Intro: risk sharing

Our paper introduces three innovations with respect to the standard literature. First, labor supply is explicitly recognized and modeled as an endogenous variable, that responds to exogenous shocks. Actually, one of the main topic of the paper is precisely to investigate how adverse income shocks trigger changes in labor supply at the household, and possibly at the risk sharing group level. Second, we consider variations in non labor income, but also in wage. While price uncertainty arguably plays an important role in real life (if only because wage variations are a crucial component of income shocks), not much is known about optimal risk sharing in this context. In the paper, we provide an exhaustive, theoretical characterization of efficient risk sharing contracts in a general context of uncertainty on wages and incomes; as well as an empirical implementation. Finally, our setting is fairly general; in particular, it allows for different levels of risk aversion, both within and between households. We offer a detailed discussion of identification issues in this framework; we show, in particular, that preferences and the decision process are non parametrically identifiable,[1] and we show how long panels can be used in practice to achieve identification.

Related literature

Outline

*Townsend acknowledges support from the NICHD-NSF.

[1]For the distinction between identification and identifiability, see Chiappori and Ekeland (2008).

2 Theory

2.1 Sharing wage risk: an introductory example

The risks linked to price fluctuations are less studied in the theoretical literature than those affecting income. Still, they raise interesting issues. One is that agents can respond to price (or wage) variations by adjusting their demand (or labor supply) behavior in an optimal way. The maximization implicit in this process, in turn, introduces an element of convexity into the picture - remember that max is a convex operator. Therefore, the patterns of risk sharing agreements, when the risks are linked with price variations, are somehow specific.

The remainder of this section provides a theoretical approach to the general problem. Here, we illustrate the issue with a simple example. Consider a two agent household, with two commodities - one agent's leisure, consumed by that agent only, and a private consumption good consumed by both. Assume, moreover, that agent 1 is risk neutral and consumes only the consumption good, while agent 2, who also consumes leisure, is risk averse (with respect to income shocks). Formally, using Cobb-Douglas preferences::

$$U^1\left(C^1\right) = C^1 \text{ and } U^2\left(C^2, L^2\right) = \frac{\left(L^2 C^2\right)^{1-\gamma}}{1-\gamma}$$

with $1/2 \leq \gamma$. Finally, the household faces a linear budget constraint; let w_2 denote 2's wages, and y (total) non labor income.

Since agent 1 is risk neutral, one may expect that she will bear all the risk. However, in the presence of wage fluctuations, it is *not* the case that agent 2's consumption, labor supply or even utility will remain constant. To see why, note first that ex ante efficiency implies ex post efficiency, and that the latter has a simple translation. Consider the household as a small economy in which all commodities are privately consumed. Any efficient allocation can be decentralized via adequate transfers of the consumption good (the numeraire). The decentralization process is simple; first, split non labor income y so that 2 receives ρ and 1 receives $y - \rho$; second, let 2 choose optimally her consumption and labor supply subject to the budget constraint $C^2 + w_2 L^2 = w_2 T + \rho$ (and note that 1 will consume all her share $y - \rho$). Crucially, ρ, called the 'sharing rule', is in general a function of (w_2, y); i.e., the tranfers need not be the same for all price/income bundles (see Chiappori 1992 for a precise analysis). By the first welfare theorem, any ρ function actually generates an ex post efficient allocation; however, as we shall see, ex ante efficiency strongly constrains the form of ρ.

For some given ρ, the labor supply and consumption of 2, when not hitting the constraint $L^2 = T$, have the form:

$$L^2 = \frac{\rho + w_2 T}{2w_2}, C^2 = \frac{\rho + w_2 T}{2}$$

leading to the following, indirect utility:

$$V^2(\rho, w_2) = \frac{2^{\gamma-1}}{1-\gamma}(\rho + w_2 T)^{2-2\gamma} w_2^{-(1-\gamma)}$$

which is indeed concave in ρ when $1/2 \leq \gamma$; note, however, that it is not necessarily concave in w_2.[2]

We now turn to ex ante efficiency, i.e. efficient risk sharing. By standard arguments, this requires that the ratio of the two members' marginal utilities of income remain constant. Given the risk neutrality assumption for agent 1, this boils down to the marginal utility of income of agent 2 remaining constant:

$$V^2_\rho = 2^\gamma (\rho + w_2 T)^{1-2\gamma} w_2^{-(1-\gamma)} = K$$

This gives the general form of ρ up to one constant:

$$\rho = 2K'.w_2^{\frac{1-\gamma}{1-2\gamma}} - w_2 T$$

where K' is a constant depending on the respective Pareto weights. In the end:

$$L^2 = K'.w_2^{\frac{-\gamma}{2\gamma-1}}, C^2 = K'.w_2^{-\frac{1-\gamma}{2\gamma-1}}$$

and the indirect utility is of the form:

$$V^2 = K''.w_2^{-\frac{1-\gamma}{2\gamma-1}}$$

for some constant K'', whereas agent 1's consumption (and utility) is given by the budget constraint:

$$C^1 = w_2 T - 2K'.w_2^{-\frac{1-\gamma}{2\gamma-1}} + y$$

As expected, because of 1's risk neutrality, 2 is sheltered from *non labor income risk* by his risk sharing agreement with 1: variations in y exclusively affect 1's consumption . But 2's consumption, labor supply and welfare all fluctuate with his wage: sheltering 2 from fluctuations in his own wage would be feasible but inefficient.

[2] The corresponsing index of relative aversion, computed using the total potential income $\rho + Tw_2$, is $2\gamma - 1$; it is positive if and only if $\gamma \geq 1/2$.

2.2 The general framework

We consider a risk sharing group consisting of H households. Household h, $h = 1, ..., H$ consists of I_h individual members. Commodities are individual leisure $L^{i,h}$ plus one composite good C which is privately consumed by each agent in the risk sharing group; aggregate consumption is thus $C^h = \sum_{i=1}^{I_h} C^{i,h}$ at the household level, and $C = \sum_{i,h} C^{i,h} = \sum_h C^h$ for the large community risk sharing group. Note consumption and utility vary with wage.

Individual preferences are egoistic, i.e. of the form $U^{i,h}\left(L^{i,h}, C^{i,h}\right)$strictly monotone increasing. For each household h, we consider the vector $\left(w^h, y^h\right) = \left(w^{1,h}, ..., w^{I_h,h}, y^{1,h}, ..., y^{I_h,h}, \tau_s^h\right)$ of individual wages and non labor incomes within household h; note that non labor incomes include remittances, and more generally all transfers received from outside the risk sharing group. This vector fluctuates randomly, following some known distribution; in what follows, everything should be understood as conditional on the distributions. Specifically, there are S possible states of the world; let $\left(w_s^h, y_s^h\right) = \left(w_s^{1,h}, ..., w_s^{I_h,h}, y_s^{1,h}, ..., y_s^{I_h,h}, \tau_s^h\right), s = 1, ..., S$ denote the realization of the individual wages and incomes vector and let τ_s^h denote the transfer received by the household in state s, which is reached with common agreed upon probability π_s. We denote $L_x^{i,h}$ and $C_s^{i,h}$ the leisure and consumption of member i of household h in state s.

We are interested in efficient risk sharing, both within and across households.

2.3 The household level

We start with the analysis of risk sharing within a given household h, taking as given the transfers $\tau^h = \left(\tau_1^h, ..., \tau_S^h\right)$ received by the household in each state of the world, $s = 1, ..., S$.

2.3.1 Ex post efficiency and the sharing rule

A first characterization of efficient risk sharing within the household relies on the notion of sharing rule, which directly generalizes the concept introduced in the previous example. Specifically, efficiency requires that household h behaves as if it were solving the program:

$$\max_{L,C} \sum_{i=1}^{I_h} \mu^{i,h} \sum_s \pi_s \left[U^{i,h}\left(L_s^{i,h}, C_s^{i,h}\right)\right] \tag{1}$$

under household h overall budget constraint in each state s,

$$\sum_{i=1}^{I_h} C_s^{i,h} + \sum_{i=1}^{I_h} w_s^{i,h} L_s^{i,h} = \sum_{i=1}^{I_h} w_s^{i,h} T^i + y_s^{1,h} + ... + y_s^{I_h,h} + \tau_s^h, s = 1, ..., S \tag{2}$$

Here, the $\mu^{i,h}$ are the respective Pareto weights across individuals within household h, normalized by the convention $\sum_i \mu^{i,h} = 1$. Note that efficiency implies that the $\mu^{i,h}$ do not depend on the *realization* of wages and incomes; however, they may of course depend on their *ex ante distribution*, as assessed when the risk sharing contract was signed.

In what follows, we define Y_s^h as the sum of individual non labor incomes of household h in state s, and X_s^h as the sum of Y_s^h and the transfer τ_s^h received from other households in the risk sharing group:

$$Y_s^h = y_s^{1,h} + ... + y_s^{I_h,h}, \qquad X_s^h = Y_s^h + \tau_s^h$$

A first remark is that any ex ante efficient allocation is also ex post efficient; therefore, for each state of the world s the household solves:

$$\max_{L^{i,h},C^{i,h}} \sum_{i=1}^{I_h} \mu^{i,h} U^{i,h} \left(L_s^{i,h}, C_s^{i,h} \right) \tag{3}$$

under the budget constraint

$$\sum_{i=1}^{I_h} C_s^{i,h} + \sum_{i=1}^{I_h} w_s^{i,h} L_s^{i,h} = \sum_{i=1}^{I_h} w_s^{i,h} T^i + X_s^h \tag{4}$$

The following Lemma directly extends standard results of the collective model:

Lemma 1 *(Intra household sharing rule) Program (3) can be decentralized in the following way: for each state of the world s, there exists a vector $\rho_s^h = \left(\rho_s^{1,h}, ..., \rho_s^{I_h,h} \right)$ (the 'sharing rule'), with $\sum_i \rho_s^{i,h} = X_s^h$, such that for $i = 1, ..., I_h$ member i solves*

$$\max_{L^i,C^i} U^i \left(L_s^{i,h}, C_s^{i,h} \right) \tag{5}$$

under the budget constraint

$$C_s^{i,h} + w_s^{i,h} L_s^{i,h} = w_s^{i,h} T^i + \rho_s^{i,h}$$

Proof. *The statement follows from the second welfare theorem: considering the household as a small economy, we know that any efficient allocation can be decentralized through adequate income tranfers.* ■

The existence of a sharing rule has a very specific interpretation. In an ex post efficient group, the labor supply of any member, say i, may depend on the realization of *all* wages in the household, not only i's own wage. But all other wages matter only through the sharing rule, which summarizes the transfers taking place across members of the groups. By the same token, we shall see that labor supplies within household h depend not only on wages and incomes of individuals belonging to h, but on wages and incomes of *all* individuals within the risk sharing group. However, wages and incomes of individuals belonging to other households matter only through the transfer τ^h, which itself matters through X_s^h. At the household level, τ^h is taken as given; how it varies with wages will be analyzed at the level of the risk sharing group below.

In the previous statement, the sharing rule $\rho_s^{i,h}$ depends on the particular state of the world under consideration. Note, however, that the state of the world is totally defined by the realization of wages, non labor incomes and tranfer, i.e. by the vector $\left(w_s^h, y_s^h, \tau_s^h\right) = \left(w_s^{1,h}, ..., w_s^{I_h,h}, y_s^{1,h}, ..., y_s^{I_h,h}, \tau_s^h\right)$. In other words, the sharing rule defines a *function* from the set of such vectors, which is a subset of $\mathbb{R}^{I_h} \times \mathbb{R}^{I_h} \times \mathbb{R}$, to $\mathbb{R}^{I_h}$. We denote this function $\tilde{\rho}^h = \left(\tilde{\rho}^{1,h}, ..., \tilde{\rho}^{I_h,h}\right)$; therefore:

$$\rho_s^{i,h} = \tilde{\rho}^{i,h}\left(w_s^{1,h}, ..., w_s^{I_h,h}, y_s^{1,h}, ..., y_s^{I_h,h}, \tau_s^h\right)$$

for all $\left(w_s^{1,h}, ..., w_s^{I_h,h}, y_s^{1,h}, ..., y_s^{I_h,h}, \tau_s^h\right)$. Here, the functions $\tilde{\rho}^{i,h}$ satisfy the relationship:

$$\sum_i^{i,h} \tilde{\rho}\left(w_s^{1,h}, ..., w_s^{I_h,h}, y_s^{1,h}, ..., y_s^{I_h,h}, \tau_s^h\right) = \sum_i y_s^{i,h} + \tau_s^h$$

for all $\left(w^1, ..., w^{I_h}, y^1, ..., y^{I_h}, \tau\right) \in \mathbb{R}^{I_h} \times \mathbb{R}^{I_h} \times \mathbb{R}$. In what follows, we moreover assume that this function is continuously differentiable.

It is important to note that the existence of a sharing rule stems from *ex post* efficiency, and is completely independent from risk sharing; conversely, any household for which a sharing rule exists will be *ex post* efficient. But it may not, and in general will not, be *ex ante* efficient. Efficient risk sharing has strong implications for the sharing rule, that will be derived later on.

Let us introduce some notations at that point. The (Marshallian) demand for leisure that solves program (5) will be denotes $H^{i,h}$; note that it is a function of i's wage, $w_s^{i,h}$, and share of non labor income, $\rho_s^{i,h}$. The resulting, indirect utility is denoted $v^{i,h}$, and again is a function of $w_s^{i,h}$ and $\rho_s^{i,h}$. Both functions $H^{i,h}$ and $v^{i,h}$ only depend on i's preferences, while one of their argument - the sharing rule $\rho_s^{i,h}$ - depends on the decision process.

2.3.2 Ex ante efficiency

We now introduce additional restrictions on the sharing rule, reflecting the fact that it must implement an allocation of risk that is *ex ante* efficient. We start with a restatement of a well-known property, the so-called Mutuality Principle. Again, let X_s^h denote total non labor income available at the household level:

$$X_s^h = \sum_i y_s^{i,h} + \tau_s^h$$

Lemma 2 *(Mutuality Principle)Assume that agents are strictly risk averse. If $\tilde{\rho}^h$ implements an (ex ante) efficient allocation, and if two states s and s' are such that $w_s^h = w_{s'}^h$ and $X_s^h = X_{s'}^h$ then*

$$\tilde{\rho}^{i,h}\left(w_s^{1,h}, ..., w_s^{I_h,h}, y_s^{1,h}, ..., y_s^{I_h,h}, \tau_s^h\right) = \tilde{\rho}^{i,h}\left(w_{s'}^{1,h}, ..., w_{s'}^{I_h,h}, y_{s'}^{1,h}, ..., y_{s'}^{I_h,h}, \tau_{s'}^h\right)$$

for all i. Therefore there exists functions $\rho^{i,h}, i = 1, ..., I_h$ from $\mathbb{R}^{I_h} \times \mathbb{R}$ to $\mathbb{R}^{I_h}$ such that

$$\tilde{\rho}^{i,h}\left(w_s^{1,h}, ..., w_s^{I_h,h}, y_s^{1,h}, ..., y_s^{I_h,h}, \tau_s^h\right) = \rho^{i,h}\left(w_s^h, X_s^h\right)$$

Moreover, we have that

$$0 < \frac{\partial \rho^{i,h}\left(w_s^h, X_s^h\right)}{\partial X_s^h} < 1$$

Proof. *Assume two states s and s' are such that $w_s^h = w_{s'}^h$ and $X_s^h = X_{s'}^h$ but $\rho_s^{i,h} \neq \rho_{s'}^{i,h}$ for at least one agent. Consider the set J of agents j such that $\rho_s^{j,h} \neq \rho_{s'}^{j,h}$, and the sharing rule $\bar{\rho}$ defined as follows:*

- $\bar{\rho}_t^{i,h} = \rho_t^{i,h}$ *for all i and all $t \neq s, s'$*
- $\bar{\rho}_s^{i,h} = \rho_s^{i,h}$ *and* $\bar{\rho}_{s'}^{i,h} = \rho_{s'}^{i,h}$ *for all $i \notin J$*
- $\bar{\rho}_s^{j,h} = \bar{\rho}_{s'}^{j,h} = \frac{1}{\pi_s + \pi_{s'}}\left(\pi_s \rho_s^{j,h} + \pi_{s'} \rho_{s'}^{j,h}\right)$*for all $i \in J$*

Then $\bar{\rho}$ satisfies the budget constraint in each states and Pareto dominates ρ ex ante, a contradiction.

Regarding the second property, note that an ex ante efficient allocation solves, for some Pareto weights $\mu^{1,h}, ..., \mu^{I_h,h}$, the program:

$$\max_{\rho^{i,h}} \sum_i \mu^{i,h} \sum_s \pi_s v^{i,h}\left(w_s^{i,h}, \rho_s^{i,h}\right) \tag{6}$$

under the budget constraint

$$\sum_i \rho_s^{i,h} = X_s^h \tag{7}$$

and where $\rho_s^{i,h} = \rho^{i,h}\left(w_s^h, X_s^h\right)$. *First order conditions give:*

$$\mu^{i,h}\pi_s \frac{\partial v^{i,h}\left(w_s^{i,h}, \rho_s^{i,h}\right)}{\partial \rho_s^{i,h}} = \lambda_s$$

where λ_s *is the Lagrange multiplier associated with the budget constraint. Therefore, for any* $i \geq 1$*:*

$$\mu^{i,h}\frac{\partial v^{i,h}\left(w_s^{i,h}, \rho_s^{i,h}\right)}{\partial \rho_s^{i,h}} = \frac{\lambda_s}{\pi_s} = \mu^{1,h}\frac{\partial v^{1,h}\left(w_s^{1,h}, \rho_s^{1,h}\right)}{\partial \rho_s^{1,h}}$$

Differentiating with respect to X_s^h*:*

$$\mu^{i,h}\frac{\partial^2 v^{i,h}\left(w_s^{i,h}, \rho_s^{i,h}\right)}{\left(\partial \rho_s^{i,h}\right)^2}\frac{\partial \rho^{i,h}}{\partial X_s^h} = \mu^{1,h}\frac{\partial^2 v^{1,h}\left(w_s^{1,h}, \rho_s^{1,h}\right)}{\left(\partial \rho_s^{1,h}\right)^2}\frac{\partial \rho^{1,h}}{\partial X_s^h}$$

and dividing by the previous equation:

$$\frac{1}{RT^{i,h}\left(w_s^{i,h}, \rho_s^{i,h}\right)}\frac{\partial \rho^{i,h}}{\partial X_s^h} = \frac{1}{RT^{1,h}\left(w_s^{1,h}, \rho_s^{1,h}\right)}\frac{\partial \rho^{1,h}}{\partial X_s^h}$$

where $RT^{i,h}\left(w_s^{i,h}, \rho_s^{i,h}\right)$ *denotes i's risk tolerance:*

$$RT^{i,h}\left(w_s^{i,h}, \rho_s^{i,h}\right) = -\frac{\partial v^{i,h}\left(w_s^{i,h}, \rho_s^{i,h}\right)/\partial \rho_s^{i,h}}{\partial^2 v^{i,h}\left(w_s^{i,h}, \rho_s^{i,h}\right)/\left(\partial \rho_s^{i,h}\right)^2}$$

which is positive by assumption. Since

$$\sum_i \rho^{i,h}\left(w_s^h, X_s^h\right) = X_s^h$$

we have that

$$\sum_i \frac{\partial \rho^{i,h}\left(w_s^h, X_s^h\right)}{\partial X_s^h} = 1$$

and finally

$$\frac{\partial \rho^{i,h}}{\partial X_s^h} = \frac{RT^{i,h}\left(w_s^{i,h}, \rho_s^{i,h}\right)}{\sum_j RT^{j,h}\left(w_s^{j,h}, \rho_s^{j,h}\right)}$$

which implies the last inequality of the Lemma. ■

Lemma (2) states two results. One is that that an ex ante efficient allocation only depends on total non labor income X_s^h, not on its various components. In other words, the decision process satisfies *income pooling* *over realizations*, in the sense that only the realization of total income X_s^h matters, not the realization of its various components. Of course, it is still the case that the *ex ante distribution* of these components may and in general will matter for the choice of a particular Pareto efficient outcome on the frontier. Efficient risk sharing does *not* mean that wealthy people have the same consumption and labor supply as poor ones - but simply that the risk arising from wealth *fluctuations* is optimally spread within the group.

The second implication of Lemma (2) is that the risk associated with fluctuations of non labor income is shared between members: fluctuations in aggregate non labor income at the sharing group level affect each agent's share, hence their consumption and labor supply. Moreover, the distribution of the marginal fluctuation (reflected by the partial of the sharing rule) reflects the agents' respective risk tolerances; more risk tolerant individuals bear a larger fraction of the marginal risk, as demonstrated in the proof.

In summary, the property of efficient risk-sharing rules *regarding income risk* are direct transposition of standard properties. However, as suggested by the introductory example, sharing wage risk is a more complex matter. One can actually show that there exist restrictions regarding the relationship between wage variations and transfers, that take the form of partial differential equations in the $\rho^{i,h}$. For brevity, we do not state these conditions here; they are available from the authors.

2.3.3 Identification: ordinal preferences

We now consider the issue of identifiability. Assume that the econometrician can observe how individual labor supplies vary in response to shocks affecting incomes and wages. Is it possible to recover the underlying structural model, i.e. individual preferences and the decision process (as summarized by the sharing rule)?

Our first claim is that, under the sole assumption of *ex post* efficiency, it is possible to recover individual (ordinal) preferences and the sharing rule, up to constants, from the observation of individual labor supplies.

Assume, therefore, that we observe realized wages and household aggregate income in each state s, $\left(w_s^h, X_s^h\right) \in \mathbb{R}^{I_h} \times \mathbb{R}$, and the resulting, *individual* labor supplies (or demands for leisure) for each agent. However, individual consumptions are not observed (only aggregate household consumption can be deduced from the budget constraint). The next Proposition is a direct transposition of a standard result of the collective literature (see Chiappori 88, 92). Formally:

Proposition 3 *(Ordinal Identifiability) Generically, the sharing rule ρ^h is identified from labor supplies up to $(n-1)$ additive constants. That* *is,if two sharing rules, ρ and $\bar{\rho}$, generate the same labor supplies, it must be the case that for all $(w, X) \in \mathbb{R}^{I_h} \times \mathbb{R}$, we have:*

$$\rho^{i,h}(w, X) = \bar{\rho}^{i,h}(w, X) + K_i$$

for all i, with $\sum_i K_i = 0$. For any choice of the constants, individual preferences are each identified up to an increasing transform, and for each choice of the transforms the Pareto weights $\mu^{i,h}$ in (3) are exactly recovered.

Proof. *Let $L^{i,h}\left(w_s^h, X_s^h\right)$ be the observed demand for leisure of individual i, as a function of the vector of wages and the household's aggregate, non labor income. From (5), we have that*

$$L^{i,h}\left(w_s^h, X_s^h\right) = H^{i,h}\left(w_s^{i,h}, \rho^{i,h}\left(w_s^h, X_s^h\right)\right)$$

where $H^{i,h}$ denotes i's Marshallian demand. Therefore:

$$\frac{\partial L^{i,h}/\partial w_s^{k,h}}{\partial L^{i,h}/\partial X_s^h} = \frac{\partial \rho^{i,h}/\partial w_s^{k,h}}{\partial \rho^{i,h}/\partial X_s^h} \text{ for all } k \neq i$$

and moreover

$$\sum_i \rho^{i,h}\left(w_s^h, X_s^h\right) = X_s^h$$

implies that

$$\sum_i \frac{\partial \rho^{i,h}}{\partial w_s^{k,h}} = 0 \quad \text{for all } k,$$

$$\sum_i \frac{\partial \rho^{i,h}}{\partial X_s^h} = 1$$

Hence the I_h functions $\left(\rho^{1,h}, ..., \rho^{I_h,h}\right)$ satisfy a system of $\left(I_h\left(I_h - 1\right) + I_h + 1 = \left(I_h\right)^2 + 1\right)$ PDEs. In the case of a two-member group, identification was proved by Chiappori (1988, 1992); for the general case ($I_h \geq 2$), it follows from a result of Chiappori and Ekeland (2008). In both cases, testable (overidentifying) restrictions can be derived. ■

In particular, for any $(w, X) \in \mathbb{R}^{I_h} \times \mathbb{R}$, where $w = \left(w^1, ..., w^{I_h}\right)$ define $V^{i,h}$ for each i as:

$$V^{i,h}(w, X) \equiv v^{i,h}\left(w^i, \rho^{i,h}(w, X)\right) \tag{8}$$

where, as above, $v^{i,h}$ is member i's indirect utility. $V^{i,h}$ is often called i's *collective indirect utility*; it defines the utility level reached by i for any realization of wages and aggregate income. Note that, unlike $v^{i,h}$ which only depends on i's preferences, $V^{i,h}$ depends on both preferences and the household's decision process (as summarized by the sharing rule ρ^h).

A result by Chiappori and Ekeland (2008) implies that these collective indirect utilities are *exactly* identified (ordinally, ie. up to an increasing transform) from the observation of individual labor supplies; while the sharing rule is identified only up to additive constants, the latter have no impact on indirect utilities, and are therefore welfare irrelevant. In other words, one can exactly identify functions $\left(\bar{V}^{i,h}(w,X)\right), i = 1,...,I_h$ such that any alternative solution $\left(V^{i,h}(w,X)\right), i = 1,...,I_h$ must be such that, for all $(w,X) \in \mathbb{R}^{I_h} \times \mathbb{R}$:

$$V^{i,h}(w,X) = F^{i,h}\left[\bar{V}^{i,h}(w,X)\right] \tag{9}$$

where $F^{i,h}$ is strictly increasing. Moreover, overidentifying restrictions are generated.

In the remaining of the paper we assume that $\frac{\partial V^{i,h}}{\partial X} > 0$.

2.3.4 Efficient risk sharing and cardinal identification

The previous result states that ex post efficiency provide conditions that allow to recover an *ordinal* representation of indirect utilities from the observation of labor supply behavior. We now consider the additional restrictions generated by ex ante efficiency. We shall see, in particular, that they are sufficient to identify *cardinal* representations as well (i.e., identify each indirect utility up to an affine transform). Moreover, additional, testable conditions are generated.

To see why this result holds, let, as before, $\rho_s^{i,h}$ denote $\rho^{i,h}\left(w_s^h, X_s^h\right)$. Using Lemma 1, we see that Program (1) can be written as:

$$\max_{\rho^{i,h}} \sum_i \mu^{i,h} \sum_s \pi_s v^{i,h}\left(w_s^{i,h}, \rho_s^{i,h}\right) \tag{10}$$

under the budget constraint

$$\sum_i \rho_s^{i,h} = X_s^h \tag{11}$$

which must hold in each state of the world. The first order conditions of program (10) give:

$$\pi_s \mu^{i,h} \frac{\partial v^{i,h}\left(w_s^{i,h}, \rho_s^{i,h}\right)}{\partial \rho_s^{i,h}} = \lambda_s \tag{12}$$

where λ_s is the Lagrange multiplier of the budget constraint in state S. In particular:

$$\frac{\partial v^i\left(w_s^{i,h}, \rho_s^{i,h}\right)/\partial \rho_s^{i,h}}{\partial v^j\left(w_s^{j,h}, \rho_s^{j,h}\right)/\partial \rho_s^{j,h}} = \frac{\mu^{j,h}}{\mu^{i,h}}$$

which is the familiar condition that the ratio of marginal utility of income for any two members does not depend on the state. Here the statement refers to marginal utility income in the indirect utility function.

Also, from (8):

$$\frac{\partial V^i\left(w_s^h, X_s^h\right)}{\partial X_s^h} = \frac{\partial v^i\left(w_s^{i,h}, \rho_s^{i,h}\right)}{\partial \rho_s^{i,h}} \frac{\partial \rho^{i,h}\left(w_s^h, X_s^h\right)}{\partial X_s^h} \tag{13}$$

therefore:

$$\frac{\partial V^i\left(w_s^h, X_s^h\right)/\partial X_s^h}{\partial V^j\left(w_s^h, X_s^h\right)/\partial X_s^h} = \frac{\mu^{j,h}}{\mu^{i,h}} \frac{\partial \rho^{i,h}\left(w_s^h, X_s^h\right)/\partial X_s^h}{\partial \rho^{j,h}\left(w_s^h, X_s^h\right)/\partial X_s^h} \tag{14}$$

where the ratio $\frac{\partial \rho^{i,h}\left(w_s^h, X_s^h\right)/\partial X_s^h}{\partial \rho^{j,h}\left(w_s^h, X_s^h\right)/\partial X_s^h}$ is known from Proposition 3.

Our identification result states that this equation is sufficient to identify the cardinal representation of each V^i. Formally:

Proposition 4 *Generically, under the efficient risk sharing assumption, a cardinal representation of the V^i is exactly identifiable from individual labor supplies. Moreover, overidentifying restrictions are generated.*

Proof. *Take a particular solution $\bar{V}^1, ..., \bar{V}^{I_h}$, then we know from (9) that any other solution must take the form*

$$V^i(w, X) = F^i\left[\bar{V}^i(w, X)\right], i = 1, ..., I_h$$

for all $(w, X) \in \mathbb{R}^{I_h} \times \mathbb{R}$. Applying (14) for $j = 1$:

$$\frac{\partial V^i\left(w_s^h, X_s^h\right)/\partial X}{\partial V^1\left(w_s^h, X_s^h\right)/\partial X} = \frac{dF^i/dV}{dF^1/dV} \times \frac{\partial \bar{V}^i\left(w_s^h, X_s^h\right)/\partial X^h}{\partial \bar{V}^1\left(w_s^h, X_s^h\right)/\partial X^h} = K_{i,h} \frac{\partial \rho^{i,h}\left(w_s^h, X_s^h\right)/\partial X_s^h}{\partial \rho^{1,h}\left(w_s^h, X_s^h\right)/\partial X_s^h}$$

where $K_{i,h} = \frac{\mu^{1,h}}{\mu^{i,h}}$. Hence

$$\frac{dF^i}{dV}\left[\bar{V}^i\left(w_s^h, X_s^h\right)\right] = K_{i,h} \frac{\partial \rho^{i,h}\left(w_s^h, X_s^h\right)/\partial X}{\partial \rho^{1,h}\left(w_s^h, X_s^h\right)/\partial X} \frac{\partial \bar{V}^1\left(w_s^h, X_s^h\right)/\partial X_s^h}{\partial \bar{V}^i\left(w_s^h, X_s^h\right)/\partial X_s^h} \frac{dF^1}{dV}\left[\bar{V}^1\left(w_s^h, X_s^h\right)\right] \tag{15}$$

We now proceed in two steps. First, we show that if F^1 is chosen arbitrarily, then these equations exactly pin down the cardinal representations of F^2,...,F^{I_h}, and that additional restrictions are generated. Second, we show that these additional restrictions identify F^1 up to an affine transform

Step 1: *Start with an arbitrary F^1, and consider (15) as an equation in F^i. Since $\frac{\partial V^i}{\partial X} > 0$, the change in variables ϕ:*

$$\left(w_s^{1,h}, ..., w_s^{I_h,h}, X_s^h\right) \to \left(w_s^{1,h}, ..., w_s^{I_h,h}, \bar{V}^i\right)$$

is invertible; in particular, we may define the function Ξ^i *as the inverse of* $\bar{V}^i$ *by:*

$$\bar{V}^i(w,X)=v\Leftrightarrow X=\Xi^i(w,v)$$

for all $(w,X)\in\mathbb{R}^{I_h}\times\mathbb{R}$*. Define the functions* A_i *as:*

$$A_i(w,X)=K_{i,1}\frac{\partial\rho^i(w,X)/\partial X}{\partial\rho^1(w,X)/\partial X}\,\frac{\partial\bar{V}^1(w,X)/\partial X}{\partial\bar{V}^i(w,X)/\partial X}\frac{dF^1}{dV}\left[\bar{V}^1(w,X)\right]$$

for all $(w,X)\in\mathbb{R}^{I_h}\times\mathbb{R}$*, then (15) becomes*

$$\frac{dF^i}{dV}\left(\bar{V}^i\right)=A_i\left(w,\Xi^i\left(w,\bar{V}^i\right)\right)\tag{16}$$

for all $(w,X)\in\mathbb{R}^{I_h}\times\mathbb{R}$*. We conclude that once* F^1 *is known, the derivative of* F^i *is identified up to the multiplicative constant* $K_{i,1}$*, so that* F^i *itself is identified up to an affine transform (i.e., cardinally identified).*

Step 2 *We now exploit a last restriction, i.e. that the right hand side in (16) must depend only on* $\bar{V}^i$*. Define, for all* $(w,X)\in\mathbb{R}^{I_h}\times\mathbb{R}$*:*

$$B_i(w,X)=K_{i,1}\frac{\partial\rho^i(w,X)/\partial X}{\partial\rho^1(w,X)/\partial X}\,\frac{\partial\bar{V}^1(w,X)/\partial X}{\partial\bar{V}^i(w,X)/\partial X}$$

Note that the function B_i *is known, in the sense that it can be recovered from labor supply behavior. Then (15) becomes:*

$$\frac{dF^i}{dV}\left[\bar{V}^i\right]=B_i\left(w,\Xi^i\left(w^i,\bar{V}^i\right)\right)\,\frac{dF^1}{dV}\left[\bar{V}^1\left(w,\Xi^i\left(w^i,\bar{V}^i\right)\right)\right]\tag{17}$$

and the condition that the right hand side only depends on $\bar{V}^i$ *becomes for any* $j\neq i$*:*

$$\left(\frac{\partial B_i}{\partial w^j}+\frac{\partial B_i}{\partial X}\frac{\partial\Xi^i}{\partial\bar{V}^i}\frac{\partial\bar{V}^i}{\partial w^j}\right)\frac{dF^1}{dV}+B_i\,\frac{d^2F^1}{(dV)^2}\left(\frac{\partial\bar{V}^1}{\partial w^j}+\frac{\partial\bar{V}^1}{\partial X}\frac{\partial\Xi^i}{\partial\bar{V}^i}\frac{\partial\bar{V}^i}{\partial w^j}\right)=0,$$

or equivalently, since $\frac{dF^1}{dV}>0$*:*

$$\frac{d^2F^1/(dV)^2}{dF^1/dV}=-\frac{1}{B_i}\frac{\frac{\partial B_i}{\partial w^j}+\frac{\partial B_i}{\partial X}\frac{\partial\Xi^i}{\partial\bar{V}^i}\frac{\partial\bar{V}^i}{\partial w^j}}{\frac{\partial\bar{V}^1}{\partial w^j}+\frac{\partial\bar{V}^1}{\partial X}\frac{\partial\Xi^i}{\partial\bar{V}^i}\frac{\partial\bar{V}^i}{\partial w^j}}\tag{18}$$

This equation, in turn, defines F^1 *up to an affine transform. Additional overidentifying restrictions are generated. First, the previous characterization should not depend on* j*, which requires that for all* $j,k\neq i$*:*

$$\frac{\frac{\partial B_i}{\partial w^j}+\frac{\partial B_i}{\partial X}\frac{\partial\Xi^i}{\partial\bar{V}^i}\frac{\partial\bar{V}^i}{\partial w^j}}{\frac{\partial\bar{V}^1}{\partial w^j}+\frac{\partial\bar{V}^1}{\partial X}\frac{\partial\Xi^i}{\partial\bar{V}^i}\frac{\partial\bar{V}^i}{\partial w^j}}=\frac{\frac{\partial B_i}{\partial w^k}+\frac{\partial B_i}{\partial X}\frac{\partial\Xi^i}{\partial\bar{V}^i}\frac{\partial\bar{V}^i}{\partial w^k}}{\frac{\partial\bar{V}^1}{\partial w^k}+\frac{\partial\bar{V}^1}{\partial X}\frac{\partial\Xi^i}{\partial\bar{V}^i}\frac{\partial\bar{V}^i}{\partial w^k}}$$

Second, again the right hand side of (18) should depend only on $\bar{V}^1$*.* ■

In summary, the analysis of individual labor supplies at the household level allows us to identify individual preferences and intrahousehold redistribution up to constants, as well as to test the model. These constants are welfare irrelevant, in the sense that indirect utilities are exactly recovered. Finally, the assumption of efficient intrahousehold risk sharing generates additional tests and allows to exactly recover indirect VNM utilities.

It is important to note that this result is in sharp contrast with the standard literature, which usually concludes that cardinal utilities can only be recovered up to a common increasing transform; i.e., one agent's cardinal utility can be arbitrarily chosen, and then the others are exactly pinned down. The difference stems from the fact that the standard literature only considers consumption of one commodity, whereas our model has two (consumption and labor supply) and allows the relative price (here the wage) to vary. In other words, our stronger result comes from the fact that our approach considers both income and price risks.

2.3.5 The household indirect utility

Finally, we may define the indirect utility of household h in state s as

$$\omega^h\left(w_s^h, X_s^h\right) = \sum_i \mu^{i,h} v^{i,h}\left(w_s^{i,h}, \rho^{i,h}\left(w_s^h, X_s^h\right)\right) \tag{19}$$

$$= \sum_i \mu^{i,h} V^{i,h}\left(w_s^h, X_s^h\right) \tag{20}$$

If we single out the transfers by defining

$$Z_s^h = y_s^{1,h} + ... + y_s^{I_h,h}$$

so that

$$X_s^h = Z_s^h + \tau_s^h$$

From (19), with this substitution, we know that:

$$\frac{\partial \omega^h}{\partial \tau_s}\left(w_s^h, Z_s^h + \tau_s^h\right) = \sum_i \mu^{i,h} \frac{\partial v^{i,h}}{\partial \rho^{i,h}} \frac{\partial \rho^{i,h}}{\partial X_s^h} \tag{21}$$

and (12) and (21) then imply:

$$\pi_s \frac{\partial \omega^h}{\partial \tau_s}\left(w_s^h, Z_s^h + \tau_s^h\right) = \sum_i \pi_s \mu^{i,h} \frac{\partial v^{i,h}\left(w_s^{i,h}, \rho_s^{i,h}\right)}{\partial \rho^{i,h}} \frac{\partial \rho^{i,h}}{\partial X_s^h}$$

$$= \lambda_s \left(\sum_i \frac{\partial \rho^{i,h}}{\partial X_s^h}\right) = \lambda_s \tag{22}$$

since again the budget constraint (11) implies that:

$$\sum_i \frac{\partial \rho^{i,h}}{\partial X_s^h} = 1.$$

2.4 The risk sharing group level

At the risk sharing group level, the general program can be stated as the determination of state-contingent transfers that maximize a weighted sum of utilities. Hence we have the program

$$\max_{\tau_1^1,\dots,\tau_S^H} \sum_h M^h \sum_s \pi_s \omega^h \left(w_s^h, Z_s^h + \tau_s^h \right) \tag{23}$$

under the resource constraint

$$\sum_h \tau_s^h = 0 \ \text{ for all } s \tag{24}$$

for all s. Here, M^h is the Pareto weight of household h within the risk sharing group, with the convention $\sum_h M^h = 1$. If τ is the vector $\left(\tau_1^1, ..., \tau_S^H\right)$, we have the following properties:

Proposition 5 *Efficient risk sharing implies the following properties:*

- *τ is such that, for any h and any s, the sum $Y_s^h = Z_s^h + \tau_s^h$ is a function of $Z_s = Z_s^1 + ... + Z_s^H$ only; i.e., if two states s and s' are such that $Z_s = Z_{s'}$ then $Z_s^h + \tau_s^h = Z_{s'}^h + \tau_{s'}^h$ for all h.*

- *If all households are strictly risk averse, then for all h*

$$0 < \frac{\partial Y_s^h}{\partial Z_s} < 1$$

Proof. *The first statement is also known as the Mutuality Principle. To show it, just note that the previous program can equivalently be written as:*

$$\max_{\tau_1^1,\dots,\tau_S^H} \sum_h M^h \sum_s \pi_s \omega^h \left(w_s^h, Y_s^h \right)$$

under the constraint

$$\sum_h Y_s^h = \sum_h Z_s^h \ \text{ for all } s \tag{25}$$

Since the program only depends on Z_s, so does its solution. The second result is standard. ■

Moreover, the previous identification results provide additional predictions. First order conditions for program (23) can be written as:

$$M^h \pi_s \frac{\partial \omega^h}{\partial \tau_s^h} = \Lambda_s$$

where Λ_s is the Lagrange multiplier of the resource constraint (24) in state s. Therefore, from (22):

$$M^h \lambda_s^h = \Lambda_s$$

In particular, if i (resp. j) belongs to household h (resp. k) from (10) again:

$$\frac{\partial v^{h,i}\left(w_s^i, \rho_s^{h,i}\right)/\partial \rho_s^{h,i}}{\partial v^{k,j}\left(w_s^j, \rho_s^{k,j}\right)/\partial \rho_s^{k,j}} = \frac{\lambda_s^h}{\lambda_s^k}\frac{\mu^{j,k}}{\mu^{i,h}} = \frac{M^k \mu^{j,k}}{M^h \mu^{i,h}}$$

which is the standard necessary condition for efficiency.

In words: efficient risk sharing can be performed in a two-stage manner, with transfers between households and between members of a given household; the Pareto weight of any individual is the product of the Pareto weight of the person's household, M^h, by the person's Pareto weight within her household, $\mu^{i,h}$.

3 A parametric model

We now specialize the general results derived above to a convenient functional form. We consider the Cobb-Douglas utility functions used in many applied works (see f.i. Altug-Miller 1990, Hayashi et al. 1996):

$$U^{i,h}\left(C^{i,h}, L^{i,h}\right) = \frac{\left(\left(L^{i,h}\right)^{\alpha_{i,h}}\left(C^{i,h}\right)^{1-\alpha_{i,h}}\right)^{1-\gamma_{i,h}}}{1-\gamma_{i,h}}$$

3.1 Frish demand for leisure

Demand for leisure The corresponding demand for leisure can readily be derived as a function of wages and the sharing rule::

$$L_s^{i,h} = \min\left(T, \alpha_{i,h}\frac{T w_s^{i,h} + \rho_s^{i,h}}{w_s^{i,h}}\right)$$

where T is the maximum time available. Then:

- if

$$T > \alpha_{i,h}\frac{w_s^{i,h} T + \rho_s^{i,h}}{w_s^{i,h}}, \text{ or equivalently } w_s^{i,h} > \frac{\alpha_{i,h}}{(1-\alpha_{i,h})T}\rho_s^{i,h} \quad (26)$$

then the agent works, and

$$L_s^{i,h} = \alpha_{i,h} \frac{w_s^{i,h} T + \rho_s^{i,h}}{w_s^{i,h}}$$
$$C_s^{i,h} = (1 - \alpha_{i,h}) \left(w_s^{i,h} T + \rho_s^{i,h} \right)$$

therefore, for $W =$ working

$$V_W^{i,h} = \left[K^{i,h} \left(w_s^{i,h} \right)^{-\alpha_{i,h}} \right]^{1-\gamma_{i,h}} \frac{\left(w_s^{i,h} T + \rho_s^{i,h} \right)^{1-\gamma_{i,h}}}{1 - \gamma_{i,h}}$$

where

$$K^{i,h} = \left((\alpha_{i,h})^{\alpha_{i,h}} (1 - \alpha_{i,h})^{1-\alpha_{i,h}} \right)$$

Note that

$$\frac{dV_W^{i,h}}{d\rho_s^{i,h}} = \left[K^{i,h} \left(w_s^{i,h} \right)^{-\alpha_{i,h}} \right]^{1-\gamma_{i,h}} \left(w_s^{i,h} T + \rho_s^{i,h} \right)^{-\gamma_{i,h}}$$

- while if

$$w_s^{i,h} \leq \frac{\alpha_{i,h}}{(1 - \alpha_{i,h}) T} \rho_s^{i,h}$$

then the agent does not work, and

$$L_s^{i,h} = T$$
$$C_s^{i,h} = \rho_s^{i,h}$$

so that her utility for is, for not-working, NW:

$$V_{NW}^{i,h} = \frac{\left(T^{\alpha_{i,h}} \rho_s^{i,h(1-\alpha_{i,h})} \right)^{(1-\gamma_{i,h})}}{1 - \gamma_{i,h}}$$

In summary, the indirect utility is thus:

$$v^{i,h} \left(w_s^{i,h}, \rho_s^{i,h} \right) = \frac{\left(T^{\alpha_{i,h}} \rho_s^{i,h(1-\alpha_{i,h})} \right)^{(1-\gamma_{i,h})}}{1 - \gamma_{i,h}} \quad \text{if } w_s^{i,h} T \leq \frac{\alpha_{i,h}}{1 - \alpha_{i,h}} \rho_s^{i,h} \tag{27}$$

$$= \frac{\left[K^{i,h} \left(w_s^{i,h} \right)^{-\alpha_{i,h}} \left(w_s^{i,h} T + \rho_s^{i,h} \right) \right]^{1-\gamma_{i,h}}}{1 - \gamma_{i,h}} \quad \text{if } w_s^{i,h} T \tag{28}$$

$$> \frac{\alpha_{i,h}}{1 - \alpha_{i,h}} \rho_s^{i,h} \tag{29}$$

The agent has a either a CARA (not working) or a HARA (working) utility function with respect to income risk, with a coefficient equal to $\gamma_{i,h}$. Note that utility is differentiable at the reservation wage $w_s^{i,h} = \left(\frac{1}{T}\right)\left(\frac{\alpha_{i,h}}{1-\alpha_{i,h}T}\rho_s^{i,h}\right)$, and strictly concave in ρ throughout.

An important remark, here, is the following. Compute the partial of $v^{i.h}$ in $\rho_s^{i,h}$ at the participation threshold (i.e., for $w_s^{i,h}T = \frac{\alpha_{i,h}}{1-\alpha_{i,h}}\rho_s^{i,h}$); one can readily check that it is equal to

$$(1-\alpha_{i,h})^{(1-\alpha_{i,h})(1-\gamma_{i,h})}\left(\frac{w_s^{i,h}T}{\alpha_{i,h}}\right)^{(1-\alpha_{i,h})(1-\gamma_{i,h})-1}$$

Now, by strict concavity of $v^{i,h}$, we know that this partial is strictly decreasing in $\rho_s^{i,h}$. It follows that there are two equivalent conditions for working:

$$\frac{\partial v^{i,h}\left(w_s^{i,h},\rho_s^{i,h}\right)}{\partial\rho_s^{i,h}} \geq (1-\alpha_{i,h})^{(1-\alpha_{i,h})(1-\gamma_{i,h})}\left(\frac{w_s^{i,h}T}{\alpha_{i,h}}\right)^{(1-\alpha_{i,h})(1-\gamma_{i,h})-1} \tag{30}$$

$$\text{if and only if}$$

$$\rho_s^{i,h} \leq \frac{1-\alpha_{i,h}}{\alpha_{i,h}}w_s^{i,h}T$$

and the participation condition ($\rho_s^{i,h} \leq \frac{1-\alpha_{i,h}}{\alpha_{i,h}}w_s^{i,h}T$) can equivalently be expressed by (30), which only depends on the partial of $v^{i,h}$.

To see why this characterization is important, remember that the necessary and sufficient condition for efficient risk sharing, (12), is:

$$\frac{\partial v^{i,h}\left(w_s^{i,h},\rho_s^{i,h}\right)}{\partial\rho_s^{i,h}} = \frac{\lambda_s^h}{\pi_s\mu^{i,h}} \tag{31}$$

Therefore the participation decision can be described as follows:

- if

  $$\frac{\lambda_s}{\pi_s\mu^{i,h}} \leq (1-\alpha_{i,h})^{(1-\alpha_{i,h})(1-\gamma_{i,h})}\left(\frac{w_s^{i,h}T}{\alpha_{i,h}}\right)^{(1-\alpha_{i,h})(1-\gamma_{i,h})-1}$$

  then

  $$\rho_s^{i,h} \geq \frac{1-\alpha_{i,h}}{\alpha_{i,h}}w_s^{i,h}T$$

  and the **agent does not work**. Her indirect utility is given by (27), therefore (31) becomes:

  $$(1-\alpha_{i,h})\,T^{\alpha_{i,h}(1-\gamma_{i,h})}\rho_s^{i,h(1-\alpha_{i,h})(1-\gamma_{i,h})-1} = \frac{\lambda_s^h}{\pi_s\mu^{i,h}} \tag{32}$$

which allows us to derive the sharing rule

$$\rho_s^{i,h} = \left(\frac{\lambda_s^h}{\pi_s \mu^{i,h} (1 - \alpha_{i,h})} T^{\alpha_{i,h}(1-\gamma_{i,h})} \right)^{\frac{1}{(1-\alpha_{i,h})(1-\gamma_{i,h})-1}} \tag{33}$$

and the corresponding consumption:

$$C_s^{i,h} = \rho_s^{i,h} = \left(\frac{\lambda_s^h}{\pi_s \mu^{i,h} (1 - \alpha_{i,h})} T^{\alpha_{i,h}(1-\gamma_{i,h})} \right)^{\frac{1}{(1-\alpha_{i,h})(1-\gamma_{i,h})-1}} \tag{34}$$

- if, on the contrary,

$$\frac{\lambda_s}{\pi_s \mu^{i,h}} > (1 - \alpha_{i,h})^{(1-\alpha_{i,h})(1-\gamma_{i,h})} \left(\frac{w_s^{i,h} T}{\alpha_{i,h}} \right)^{(1-\alpha_{i,h})(1-\gamma_{i,h})-1}$$

then the agent works. Her indirect utility is given by (28) for working, therefore the condition is

$$\pi_s \mu^{i,h} \left(w_s^{i,h}\right)^{-\alpha_{i,h}(1-\gamma_{i,h})} \left(T w_s^{i,h} + \rho_s^{i,h}\right)^{-\gamma_{i,h}} \left(K^{i,h}\right)^{1-\gamma_{i,h}} = \lambda_s^h \tag{35}$$

hence the sharing rule

$$\rho_s^{i,h} = \left(\frac{\lambda_s^h}{\pi_s \mu^{i,h} \left(K^{i,h}\right)^{1-\gamma_{i,h}}} \right)^{-1/\gamma_{i,h}} \left(w_s^{i,h}\right)^{-\frac{\alpha_{i,h}(1-\gamma_{i,h})}{\gamma_{i,h}}} - T w_s^{i,h} \tag{36}$$

and the Frish demand for leisure:

$$L_s^{i,h} = \alpha_{i,h} \left(\frac{\lambda_s^h}{\pi_s \mu^{i,h} \left(K^{i,h}\right)^{1-\gamma_{i,h}}} \right)^{-1/\gamma_{i,h}} \left(w_s^{i,h}\right)^{-\frac{\alpha_{i,h}+\gamma_{i,h}-\alpha_{i,h}\gamma_{i,h}}{\gamma_{i,h}}}$$

Similarly for consumption:

$$\begin{aligned} C_s^{i,h} &= (1 - \alpha_{i,h}) \left(T w_s^{i,h} + \rho_s^{i,h}\right) \\ &= (1 - \alpha_{i,h}) \left(\frac{\lambda_s^h}{\pi_s \mu^{i,h} \left(K^{i,h}\right)^{1-\gamma_{i,h}}} \right)^{-1/\gamma_{i,h}} \left(w_s^{i,h}\right)^{-\frac{\alpha_{i,h}(1-\gamma_{i,h})}{\gamma_{i,h}}} \\ &= \frac{1 - \alpha_{i,h}}{\alpha_{i,h}} w_s^{i,h} L_s^{i,h} \end{aligned} \tag{37}$$

In summary, an agent is less likely to participate, and works less hours if she participates, when:

- her wage is low

- the household is doing well, in the sense that the budget constraint is not too tight (the Lagrange multiplier λ_s^h is small)
- her Pareto weight is large: a higher status buys additional leisure.

Conversely, she works more, along the intensive and extensive margins, when either work is attractive because the wage is high, or the household has been hit by a negative shock, generating a high value for λ_s^h, and especially when her 'power' (as measured by her Pareto weight) is relatively low.

Labor supply equations Equivalently, one may define a *Frish reservation wage*, as a function of household marginal utility of income λ_s^h. This can be done, for instance, by imposing that the Frish demand for leisure just computed equals total time available T; or, equivalently, by plugging the corresponding sharing rule into the standard participation equation (26). One gets that the agent participates if her wage exceeds:

$$\bar{w}^{i,h} = \frac{\alpha_{i,h}}{T}\left(\frac{\lambda_s}{\pi_s}\right)^{\frac{1}{\left(1-\alpha_{i,h}\right)\left(1-\gamma_{i,h}\right)-1}}\left(\mu^{i,h}\right)^{-\frac{1}{\left(1-\alpha_{i,h}\right)\left(1-\gamma_{i,h}\right)-1}}\left(1-\alpha_{i,h}\right)^{-\frac{\left(1-\alpha_{i,h}\right)\left(1-\gamma_{i,h}\right)}{\left(1-\alpha_{i,h}\right)\left(1-\gamma_{i,h}\right)-1}}$$

or equivalently

$$\begin{aligned}\log \bar{w}^{i,h} = &\log \alpha_{i,h} - \log T + \frac{1}{1-\left(1-\alpha_{i,h}\right)\left(1-\gamma_{i,h}\right)}\log\left(\mu^{i,h}\right)\\ &+\frac{\left(1-\alpha_{i,h}\right)\left(1-\gamma_{i,h}\right)}{1-\left(1-\alpha_{i,h}\right)\left(1-\gamma_{i,h}\right)}\log\left(1-\alpha_{i,h}\right) - \frac{1}{1-\left(1-\alpha_{i,h}\right)\left(1-\gamma_{i,h}\right)}\log\left(\frac{\lambda_s}{\pi_s}\right)\end{aligned}$$

Conditional on participation, the demand for leisure is:

$$\begin{aligned}\log L_s^{i,h} = &\log \alpha_{i,h} + \frac{1}{\gamma_{i,h}}\log \mu^{i,h} + \frac{\left(1-\gamma_{i,h}\right)}{\gamma_{i,h}}\log\left(\left(\alpha_{i,h}\right)^{\alpha_{i,h}}\left(1-\alpha_{i,h}\right)^{1-\alpha_{i,h}}\right)\\ &-\frac{1}{\gamma_{i,h}}\log\left(\frac{\lambda_s^h}{\pi_s}\right) - \frac{1-\left(1-\alpha_{i,h}\right)\left(1-\gamma_{i,h}\right)}{\gamma_{i,h}}\log w_s^{i,h} \qquad (38)\end{aligned}$$

Moreover, at the risk sharing group level we know that:

$$M^h \lambda_s^h = \Lambda_s$$

so the reservation wage is:

$$\begin{aligned}\log \bar{w}^{i,h} = &\log \alpha_{i,h} - \log T + \frac{1}{1-\left(1-\alpha_{i,h}\right)\left(1-\gamma_{i,h}\right)}\log\left(M^h\mu^{i,h}\right)\\ &+\frac{\left(1-\alpha_{i,h}\right)\left(1-\gamma_{i,h}\right)}{1-\left(1-\alpha_{i,h}\right)\left(1-\gamma_{i,h}\right)}\log\left(1-\alpha_{i,h}\right) - \frac{1}{1-\left(1-\alpha_{i,h}\right)\left(1-\gamma_{i,h}\right)}\log\left(\frac{\Lambda_s}{\pi_s}\right) \qquad (39)\end{aligned}$$

and the final labor supply equation for member i in household h is:

$$\log L_s^{i,h} = \left(\log \alpha_{i,h} + \frac{1}{\gamma_{i,h}} \log M^h \mu^{i,h}\right) + \frac{(1-\gamma_{i,h})}{\gamma_{i,h}} \log\left((\alpha_{i,h})^{\alpha_{i,h}} (1-\alpha_{i,h})^{1-\alpha_{i,h}}\right)$$
$$-\frac{1}{\gamma_{i,h}} \log\left(\frac{\Lambda_s}{\pi_s}\right) + \frac{(1-\alpha_{i,h})(1-\gamma_{i,h})-1}{\gamma_{i,h}} \log w_s^{i,h} \quad (40)$$

where $M^h \mu^{i,h}$ is i's Pareto weight within the risk sharing group.

3.2 Practical implementation

Estimation An obvious problem with equations (39) and (40) is that neither the Pareto weights, nor the marginal utility of income Λ are observed. The idea, therefore, is to exploit the specific structure of these equations in terms of variations within and across households. We now show that the panel nature of the data allows to recover the fundamentals of the models.

Specifically, if l_s^i denotes the number of hours, equation (39) is of the form:

$$\log \bar{w}_s^{i,h} = B_{i,h} + G_{i,h} D_s \quad (41)$$

and equation (40) can similarly be written as:

$$\log\left(T - l_s^i\right) = A_{i,h} + F_{i,h} D_s - E_{i,h} \log w_s^i \quad (42)$$

Here, $w_s^{i,h}$ denotes the realization of i's *real* wage and:

$$A_{i,h} = \left(\log \alpha_{i,h} + \frac{1}{\gamma_{i,h}} \log M^h \mu^{i,h}\right) + \frac{(1-\gamma_{i,h})}{\gamma_{i,h}} \log\left((\alpha_{i,h})^{\alpha_{i,h}} (1-\alpha_{i,h})^{1-\alpha_{i,h}}\right)$$

$$B_{i,h} = \log \alpha_{i,h} - \log T + \frac{1}{1-(1-\alpha_{i,h})(1-\gamma_{i,h})} \log\left(M^h \mu^{i,h}\right)$$
$$+\frac{(1-\alpha_{i,h})(1-\gamma_{i,h})}{1-(1-\alpha_{i,h})(1-\gamma_{i,h})} \log(1-\alpha_{i,h})$$

$$G_{i,h} = \frac{1}{1-(1-\alpha_{i,h})(1-\gamma_{i,h})}$$

$$F_{i,h} = \frac{1}{\gamma_{i,h}}$$

$$D_s = -\log\left(\frac{\Lambda_s}{\pi_s}\right)$$

$$E_{i,h} = \frac{1-(1-\alpha_{i,h})(1-\gamma_{i,h})}{\gamma_{i,h}}$$

Let $\mathcal{A}$ denote the vector of parameters just defined:

$$\mathcal{A} = (A_{i,h}, B_{i,h}, G_{i,h}, E_{i,h}, F_{i,h}, D_s, i = 1, ..., I_h, h = 1, ...H, s = 1, ..., S)$$

It can be remarked that the knowledge of $\mathcal{A}$ exactly identifies the structural parameters $\left(\alpha_{i,h}, \gamma_{i,h}, M^h \mu^{i,h}\right)$ and the ratio $\frac{\Lambda_s}{\pi_s}$; indeed,

$$\begin{aligned}
\gamma_{i,h} &= \frac{1}{F_{i,h}} = \frac{1}{G_{i,h}E_{i,h}} \\
\alpha_{i,h} &= \frac{1 - G_{i,h}\gamma_{i,h}}{G_{i,h}\left(1-\gamma_{i,h}\right)} = \frac{E_{i,h}-1}{E_{i,h}G_{i,h}-1} \\
M^h \mu^{i,h} &= \exp\left[\left(A_{i,h} - \frac{\left(1-\gamma_{i,h}\right)}{\gamma_{i,h}}\log\left(\left(\alpha_{i,h}\right)^{\alpha_{i,h}}\left(1-\alpha_{i,h}\right)^{1-\alpha_{i,h}}\right) - \log\alpha_{i,h}\right)\gamma_{i,h}\right] \text{ and} \\
\frac{\Lambda_s}{\pi_s} &= \exp\left(-D_s\right)
\end{aligned} \tag{43}$$

The question, now, is whether the observation of labor supply behavior allows to *uniquely* recover the parameters $(A_{i,h}, B_{i,h}, G_{i,h}, E_{i,h}, F_{i,h}, i = 1, ..., I_h, h = 1, ...H$ and $D_s, s = 1, ..., S)$. The answer is negative; they can be recovered up to exactly two constants, as stated by the following result:

Proposition 6 *Let* $(A_{i,h}, B_{i,h}, G_{i,h}, E_{i,h}, F_{i,h}, D_s)$ *and* $\left(\bar{A}_{i,h}, \bar{B}_{i,h}, \bar{G}_{i,h}, \bar{E}_{i,h}, \bar{F}_{i,h}, \bar{D}_s\right)$ *be two vectors satisfying (41) and (42) for the same* $\bar{w}_s^{i,h}$ *and* $l_s^{i,h}$*, where* $i = 1, ..., I_h, h = 1, ...H, s = 1, ..., S$*. Then there exists two constants* $k > 0$ *and* d *such that, for all* i, h *and* s*:*

$$\begin{aligned}
A_{i,h} &= \bar{A}_{i,h} - kd\bar{F}_{i,h} \\
B_{i,h} &= \bar{B}_{i,h} - kd\bar{G}_{i,h} \\
G_{i,h} &= k\bar{G}_{i,h} \\
F_{i,h} &= k\bar{F}_{i,h} \\
D_s &= \frac{1}{k}\bar{D}_s + d \\
E_{i,h} &= \bar{E}_{i,h}
\end{aligned} \tag{44}$$

Proof. *First, the labor supply equation:*

$$\log\left(1 - l_s^i\right) = A_{i,h} + F_{i,h}D_s - E_{i,h}\log w_s^i = \bar{A}_{i,h} + \bar{F}_{i,h}\bar{D}_s - \bar{E}_{i,h}\log w_s^i$$

implies that

$$E_{i,h} = \bar{E}_{i,h}$$

and

$$A_{i,h} + F_{i,h} D_s = \bar{A}_{i,h} + \bar{F}_{i,h} \bar{D}_s$$

Similarly, participation equation requires that

$$B_{i,h} + G_{i,h} D_s = \bar{B}_{i,h} + \bar{G}_{i,h} \bar{D}_s$$

Writing these two equations in two states s and t and differencing, we get:

$$\frac{G_{i,h}}{\bar{G}_{i,h}} = \frac{\bar{D}_s - \bar{D}_t}{D_s - D_t} = \frac{F_{i,h}}{\bar{F}_{i,h}} = k$$

therefore $G_{i,h} = k\bar{G}_{i,h}, F_{i,h} = k\bar{F}_{i,h}$ and $D_s - D_t = \frac{1}{k}\left(\bar{D}_s - \bar{D}_t\right)$ which leads to

$$D_s = \frac{1}{k}\bar{D}_s + d$$

where $d = D_0 - \frac{1}{k}\bar{D}_0$. Finally, since

$$A_{i,h} + F_{i,h} D_s = A_{i,h} + k\bar{F}_{i,h}\left(\frac{1}{k}\bar{D}_s + d\right) = \bar{A}_{i,h} + \bar{F}_{i,h}\bar{D}_s \qquad (45)$$

we conclude that $A_{i,h} = \bar{A}_{i,h} - kd$, and a similar argument applies to $B_{i,h}$. ■

A consequence of this result is that in the Frish framework, the structural parameters $\left(\alpha_{i,h}, \gamma_{i,h}, M^h, \mu^{i,h}\right)$ are not uniquely identifiable from the sole observation of labor supply behavior. Indeed, individual risk aversions are identified only up to a common, multiplicative constant k. Moreover, different values of the constant generate different estimates for $\alpha_{i,h}$ and $M^h \mu^{i,h}$; for instance, replugging (44) into (43), we see that:

$$\alpha_{i,h} = \frac{E_{i,h} - 1}{E_{i,h} G_{i,h} - 1} = \frac{\bar{E}_{i,h} - 1}{k\bar{E}_{i,h}\bar{G}_{i,h} - 1}$$

To obtain full identification, we therefore need to use some additional information, namely the fact that for each state s total consumption is observed at the *risk sharing group* (although not at the *individual*) level. For any state s, let P_h denote the set of agents in household h who provide a positive labor supply, and NP_h its complement in the household. Household consumption is given by:

$$C_s^h = \sum_{i \in P_h} C_s^{i,h} + \sum_{i \in NPh} C_s^{i,h}$$

therefore, from (37) and (34):

$$C_s^h = \sum_{i \in P_h} \frac{1 - \alpha_{i,h}}{\alpha_{i,h}} w_s^{i,h} L_s^{i,h} + \sum_{i \in NPh} \left(\frac{\lambda_s^h}{\pi_s \mu^{i,h}\left(1 - \alpha_{i,h}\right)} T^{\alpha_{i,h}\left(1 - \gamma_{i,h}\right)}\right)^{\frac{1}{\left(1-\alpha_{i,h}\right)\left(1-\gamma_{i,h}\right)-1}}$$

Using the results of Proposition 6, this gives:

$$C_s^h = \sum_{i \in P_h} \frac{1-\alpha_{i,h}}{\alpha_{i,h}} w_s^{i,h} L_s^{i,h} + \sum_{i \in NP_h} e^{\bar{G}_{i,h} kd} \left[M^h \mu^{i,h} \left(1-\alpha_{i,h}\right) \right]^{\frac{-1}{\left(1-\alpha_{i,h}\right)\left(1-\gamma_{i,h}\right)-1}} \left[e^{\bar{G}_{i,h} \bar{D}_s} \right]$$

$$= \sum_{i \in P_h} \frac{1-\alpha_{i,h}}{\alpha_{i,h}} w_s^{i,h} L_s^{i,h} + \sum_{i \in NP_h} H^{i,h} . e^{\bar{G}_{i,h} \bar{D}_s}$$

where the $\bar{G}_{i,h}$ and $\bar{D}_s$ are arbitrary solutions of the labor supply estimation and

$$H^{i,h} = e^{\bar{G}_{i,h} kd} \left[M^h \mu^{i,h} \left(1-\alpha_{i,h}\right) \right]^{\frac{-1}{\left(1-\alpha_{i,h}\right)\left(1-\gamma_{i,h}\right)-1}}$$

In particular, by regressing, risk sharing group by risk sharing group, the time series of aggregate consumption over the monetary value of leisure and the terms in $e^{\bar{G}_{i,h} \bar{D}_s}$, as recovered from the labor supply equation, one gets, from the coefficient of the $w_s^{i,h} L_s^{i,h}$ term, an exact identification of the $\alpha_{i,h}$ for working agents. This, in turn, pins down exactly the constant k and generates strong overidentifying restrictions (since k should be the same for all agents); the coefficient $H^{i,h}$ then pins down the second constant d.

Remark 7 *Note, however, that we may want to estimate instead the equations:*

$$\log \bar{w}_s^{i,h} = B_{i,h} + H_s + C_{i,h} D_s$$
$$\log \left(T - l_s^{i,h}\right) = A_{i,h} + G_s + F_{i,h} D_s - E_{i,h} \log w_s^{i,h} \qquad (46)$$

where the additional term G_s and H_s capture seasonality, prices changes, etc.

Then the parameters are identified up to additional constants

4 Data

Dataset and data construction We use data from the Townsend Thai Monthly Survey. The survey is being conducted in two changwats (provinces) in semi-urban areas and in two changwats (provinces) in the rural and of Thailand. The former two provinces are Chachoengsao and Lop Buri, while the latter two are Buri Ram and Si Sa Ket. The risk sharing groups whose ID are 702, 704, 707 and 708 are in Chachoengsao, whose ID are 2702, 2710, 2713, and 2714 are in Buri Ram, whose ID are 4901, 4903, 4904, and 4906 are in Lop Buri, and whose ID are 5301, 5306, 5309, and 5310 are in Si Sa Ket. An intensive monthly survey was initiated in 1998 in 16 risk sharing groups: four risk sharing groups in each of the four provinces. The survey began with an initial risk sharing group-wide census. Every structure and every household was enumerated, and "household" units were defined based on sleeping and eating patterns. An ongoing monthly survey of approximately 45 households in each risk sharing group began in August 1998. The August 1998 survey consisted of a baseline interview on initial socioeconomic conditions of sampled households such as demographic structure, education, occupations, health, and financial situation. The monthly updates started in September 1998 and track production, financial transactions and changing socioeconomic conditions of the same households over time. All individuals, households, and residential structures in each of the 16 risk sharing groups can be identified in the monthly responses. Current version of this paper uses the dataset of 88 months.

We use the monthly observations of adult individuals whose age are between 18 and 60 years old. Observations of other individuals are excluded from the sample. Our empirical specification requires to construct hours of leisure (or equivalently number of working hours), wage rate of each individual, and full income of household. We explain how these important variables are constructed.

On number of working hours, each individual's time in hours spent on labor supply in each interview month is computed as sum of time spent on each economic activity (in hours). Economic activities includes cultivating own plots, taking care of livestock, business, fish/shrimp, and paid work[3]. However, the number of days between the two consecutive interviews differs by households. We re-adjust the computed labor supply in hours so that how many hours on average each individual supplies his labor per 30 days. Note that we consider individual's labor supply at the residence. Hence, migration is out of our consideration.

Following the past literature (Rosenzweig (1988), Townsend (1994), and Mazzocco and Saini (2007)), we assume each individual has 14 hours per day to allocate between leisure and work, and 26 days in total per month. Total time endowment per month (T) is computed as $T = 364$ based on the assumption and total time of leisure per month is computed by subtracting number of total working time per month from the total time endowment. The constructed number of working hours varies a lot between 0 and 364 hours. Especially, the distribution of positive number of working hours is close to uniform between 0.1 and 208, the latter of whom is equal to 8 hours per day. Note that we find zero working hours for 30% of all the month-adult-individual observations. So, there are non-negligible number of month-individual observations in which they do not work. Main reasons for the zero working hours are (1) elderliness, (2) intensive involvement in non-economic activities (housework or

[3] Unfortunately, number of hours spent on houseworks is not avaialbe in the data.

schooling), (3) sickness/disease and (4) seasonal fluctuation (no work in off-peak seasons). For detail, see Townsend and Yamada(2008).

We need wage rate of each individual for each month. This is straightforward if an individual is involved in a paid job as we have data on total earnings and number of working hours from the job. On the other hand, computing wage rate if an individual is involved in "non-market" economic activity such as cultivation of own fields, taking care of livestock, running own business/fishery, is demanding. Thus, we use teh following procedure to determine wage rate in each month for wach individual.

1. If we can observe a wage rate in a month, we use the wage rate for the month (no further work needed).

2. If we don't observe a wage rate in a month, but we observe a wage in another month which is not in the month of interest.: we impute the wage rate in the month of interest by using the wage rate in the another month after adjusting for age, changwat-specific time trend and monthly cycle. For the adjust ment, I just pool all wage rates and run a simple mincerian regression. I used the coefficients for the adjustment (but, this may be logically inconsistent with the selection corrected regression below. However, due to a problem of instrument variables, this is probably best I can do). To impose the following rule, I just experiment how many times each person need to show up in the wage labor market. It seems that the result does not change a lot. So I proceed with the simplest one: if a person shows up at least one time in the wage labor market, we impute his wage in other months by the following rule:

   - If we can observe wages three months prior or after the month of interest, we use the wage of nearest month for the month of interest.
   - If we cannot satisfy the condition above, we look for wages in a same calendar month but in a different year back and forth. If we can find, we use the wage in nearest year of the same calendar month.
   - If no conditions of the two above are satisfied, we look for a wage in the nearest month.
   - If no conditions above are satisfied, it means an individual has never entered wage labor market. We need other way to impute wage for these individuals.

3. In a case that a person has never show up in the wage market, we cannot utilize his information in the market. So, we need to use

cross-sectional information for the imputation. Also, a selection problem is an issue. I run the following standard selection corrected wage regression and impute wages.

5 Testing the risk sharing hypothesis

In this section, we perform an empirical test of the risk sharing hypothesis that builds on the theoretical model.

5.1 The econometric model

Let $H_s^{i,h}$ be the number of hours of work supplied by individual i in household h in state s. Empirically, a state is composed of a time period and an indicator of the unit where, according to the theory, risk sharing takes place. In the results below, s will denote a village in a given month.

We base our empirical specification on equations (26)-(27). [NOTE: Need to change the reference when merging the 2 parts of the paper] To start with, we model log-hours worked as (where $H_s^{i,h} = T - L_s^{i,h}$):

$$\log H_s^{i,h} = e_{i,h} + f_{i,h} D_s + g_{i,h} \log W_s^{i,h} + h' X_s^{i,h} + u_s^{i,h}, \qquad (47)$$

where $W_s^{i,h}$ is the hourly wage (or equivalently the marginal productivity) of the individual in state s. When the individual is not working, $W_s^{i,h}$ denotes the potential wage that she would earn on the labor market, were she participating. In addition, D_s denotes the Lagrange multiplier of the village budget constraint, and $X_s^{i,h}$ are determinants of hours that contain demographics (age dummies, household size, gender, and education dummies). We also include yearly and monthly dummies. In some specifications, we will include individual-specific fixed effects. [NOTE: I suggest to change the notation from X to R in the theory section]

Next, the equation (27) [NOTE: change reference] for the reservation wage motivates modelling labor market participation as follows:

$$P_s^{i,h} = \mathbf{1} \left\{ \log W_s^{i,h} \geq a_{i,h} + b_{i,h} D_s + c' X_s^{i,h} + d' Q_s^{i,h} + \varepsilon_s^{i,h} \right\}, \qquad (48)$$

where $P_s^{i,h} \in \{0, 1\}$ indicate whether the individual is working on the labor market or not, and where the determinants of participation include, in addition to $X_s^{i,h}$, some covariates $Q_s^{i,h}$ which do not enter wages or hours.

Exclusion restrictions are important in order to correct for sample selection bias. Given the absence of uncontroversial excluding determinants, we have experimented with several choices. We will report the results for two specifications. In the first one, participation status in the past month, and averages of participation status over months 2 to 6 and

7 to 12 prior to the interview, respectively, are used as exclusion restrictions. In the second specification, we will include lagged participation indicators of other household members. A justification for the presence of lagged participation could be the existence of sunk costs. Note that including these costs in the theoretical model would involve dynamic considerations, hence a much more complicated problem. [NOTE: should we drop that last sentence?]

We complete the model by postulating a Mincer-type equation for market log-hourly wages:

$$\log W_s^{i,h} = l_{i,h} + m' X_s^{i,h} + v_s^{i,h}, \tag{49}$$

which depends on the common determinants $X_s^{i,h}$, as well as on unobserved additive individual effects.

In the empirical analysis, it is important to account for the fact that wages may be recorded with error. Measurement error is a common feature of survey data, especially in developing countries, and our dataset is no exception. We will denote the *observed* log-wage as $\widetilde{W}_s^{i,h}$, and model it as:

$$\log \widetilde{W}_s^{i,h} = \log W_s^{i,h} + \eta_s^{i,h}, \tag{50}$$

where $\eta_s^{i,h}$ is classical measurement error. The presence of measurement error in wages causes a downward bias on the estimated labor supply elasticity ($g_{i,h}$ in equation (47)). We will adapt our econometric approach in order to address this problem.

Together, equations (47)-(50) form a joint model of hours worked, participation, true and observed hourly wages. We assume that the shocks $(\varepsilon, u, v, \eta)$ follow a stationary multivariate normal distribution conditional on regressors, with unrestricted correlation across shocks and across periods. We also assume that (ε, u, v) are independent of D_s, and of all regressors in $X_s^{i,h}$ and $Q_s^{i,h}$, and that the measurement error η is independent of all shocks and regressors. Note that u may include measurement error in hours of work.

In order to test for the null hypothesis of complete risk sharing, we will include in $X_s^{i,h}$ an household-level indicator of income shocks. Under the risk sharing hypothesis, income shocks should have no effect on participation and hours worked. We will use three variables in turn: household nonlabor income, nonlabor income including cash flows from production, and a measure of "deficits", which is computed as the sum of consumption and capital expenditures minus cash flows from production.

We will impose some restrictions on the coefficients in (47)-(50). First, we restrict the coefficients $b_{i,h}$ and $f_{i,h}$ of the Lagrange multiplier D_s to depend only on the risk-sharing unit (village or within-village network), and we include yearly dummies and seasonal dummies (monthly) in D_s. In theory, one could include a full set of interactions of individual and month dummies as controls in the hours and participation regressions. Given the large number of individuals and time periods in the data (as may be seen in Table 6), we chose a more parsimonious specification. Thus, we only imperfectly control for the amount of risk-sharing at the village or network level. This in turn implies that our test of the risk-sharing hypothesis will be conservative.

Second, we restrict the wage elasticity $g_{i,h}$ in the hours equation (47) to be constant across individuals and villages. This is a strong restriction. Allowing for individual-specific elasticities, as the theory suggests, would require better data. Indeed, estimates of *region*-specific elasticities are already very imprecise. This is because, in order to correct for measurement error, we instrument the wage in the hours equation (see below for more details on the measurement error issue). As a consequence of the instrumentation strategy, our elasticity estimates have relatively large standard errors.

Measurement error. Hourly wages are constructed as ratios of total wages divided by hours worked. As a consequence, measurement error in hours worked mechanically implies measurement error in wages, thus implying a downward bias on the wage elasticity in the labor supply equation (Borjas, 1980). To correct for measurement error, we use a second hourly wage measure which we obtain from job questionnaires, interacted with the means of payment (workers paid by the day, or by the month).[4]

Suppose to start with that all workers participate in the labor market in every period. We shall relax this assumption shortly. Then, it is easy to see that external instruments allow to recover the wage elasticity in the hours equation. To see this, let $Z_s^{i,h}$ denote the set of job form covariates. [NOTE: Z is already used as an instrument in the theory section. It is natural to use Z to denote an instrument. Would it make sense to change Z to Y (which is the sum of y's) in the theory?] We assume that Z is independent of η and u. As Z is independent of the error term $u - g\eta$ in:

$$\begin{aligned}\log H_s^{i,h} &= e_{i,h} + f_{i,h} D_s + g\left(\log \widetilde{W}_s^{i,h} - \eta_s^{i,h}\right) + h' X_s^{i,h} + u_s^{i,h} \\ &= e_{i,h} + f_{i,h} D_s + g \log \widetilde{W}_s^{i,h} + h' X_s^{i,h} + u_s^{i,h} - g\eta_s^{i,h},\end{aligned}$$

it follows that Z is a valid instrument for $\log \widetilde{W}_s^{i,h}$.

One difficulty to implement this idea is that wages, hours and instruments are only observed for labor market participants. To address

[4] In the job form questionnaire, workers paid either by the month or by the day amount to roughly 80% of all workers (remaining categories include piece rates).

this difficulty, we use a control function approach that simultaneously corrects for measurement error in wages and sample selection due to non-participation. Our approach relies on the assumption that Z (the job form wage) is independent of u *conditional* on $P = 1$ and (X, D, Q). We present our estimation strategy in the next subsection.

Self-employed. In the data, there is an additional distinction between market work (individuals working for an employer) and work outside the market (self-employed workers). We have wages for employed individuals, though not for the self-employed. Moreover, the self-employed represent a large share of active individuals in the sample (more than half, according to Table 6 below).

We assume, consistently with our neoclassical model, that workers are indifferent between working on the market or working as self-employed. Moreover, the (implicit) wage is the same in the two types of occupations (although it may depend on the individual-specific productivity through the term $l_{i,h}$). Hence, for a farmer i in period s, $W_s^{i,h}$ is the wage at which she "hires herself" from the market, and is equal to the wage she would have earned, had she chosen to work on the market.

We will present the results for the two types of participation and hours worked: participation on the wage labor market, and participation to any economic activity. Note that, for the latter, testing the risk sharing hypothesis using contemporaneous household nonlabor income suffers from a reverse causality problem, as participation and hours have a direct effect on nonlabor income. Obtaining credible estimates of the effect of nonlabor income on overall participation (as opposed to labor market participation) is therefore challenging. As an attempt to address this problem, we use instead nonlabor income lagged one month.

5.2 Estimation

To estimate the model, we treat the individual-specific effects that appear in (47)-(50), that is:

$$\{a_{i,h}, e_{i,h}, l_{i,h}, \quad h = 1, ..., N, \quad i = 1, ..., I_h\},$$

as parameters to be estimated. This "fixed-effects" approach is justified by the fact that the number of time periods is large in the data (88 months), so that each of those parameters is precisely estimated.

We now explain how we estimate the the parameters of the participation and hours equations, respectively.

Participation. We first rewrite the participation equation in reduced form, substituting (49) into (48):

$$P_s^{i,h} = \mathbf{1}\left\{l_{i,h} + m' X_s^{i,h} - a_{i,h} - b_{i,h} D_s - c' X_s^{i,h} - d' Q_s^{i,h} \geq \varepsilon_s^{i,h} - v_s^{i,h}\right\} \quad (51)$$
$$(52)$$

Let $\mathcal{X}_s^{i,h} = \left(X_s^{i,h}, D_s, Q_s^{i,h}, \delta^{i,h}\right)$ be the set of all regressors, including individual dummies $\delta^{i,h}$. Under the normality assumption, we estimate the parameters by Probit. In particular, the estimation delivers predicted probabilities:

$$\text{Prob}\left(P_s^{i,h} = 1 | \mathcal{X}_s^{i,h}\right). \quad (53)$$

Hours of work. Turning to hours, we have using (47) and substituting (50):

$$\begin{aligned}\mathbb{E}\left(\log H_s^{i,h} | P_s^{i,h} = 1, \mathcal{X}_s^{i,h}, Z_s^{i,h}\right) = & \; e_{i,h} + f_{i,h} D_s + h' X_s^{i,h} \\ & + g\mathbb{E}\left(\log \widetilde{W}_s^{i,h} | P_s^{i,h} = 1, \mathcal{X}_s^{i,h}, Z_s^{i,h}\right) \\ & + g\mathbb{E}\left(u_s^{i,h} - g\eta_s^{i,h} | P_s^{i,h} = 1, \mathcal{X}_s^{i,h}, Z_s^{i,h}\right).\end{aligned}$$

On the right-hand side of this equation, the first expectation reflects the fact that wages are contaminated by measurement error. The second expectation represents the effect of sample selection, which is due to the fact that participation to the labor market is not random. For example, some covariates in $u_s^{i,h}$ may be observed by the individual at the time she decides to participate.

Because the measurement error η is independent of all shocks and regressors we have:

$$\mathbb{E}\left(\eta_s^{i,h} | P_s^{i,h} = 1, \mathcal{X}_s^{i,h}, Z_s^{i,h}\right) = 0.$$

Moreover, as the instruments Z are assumed independent of u conditional on $P = 1$ and $\mathcal{X}$, we have:

$$\begin{aligned}\mathbb{E}\left(u_s^{i,h} | P_s^{i,h} = 1, \mathcal{X}_s^{i,h}, Z_s^{i,h}\right) = & \; \mathbb{E}\left(u_s^{i,h} | P_s^{i,h} = 1, \mathcal{X}_s^{i,h}\right) \\ = & \; \nu\lambda\left(\text{Prob}\left(P_s^{i,h} = 1 | \mathcal{X}_s^{i,h}\right)\right),\end{aligned}$$

where ν is the covariance between $u_s^{i,h}$ in (47) and $\varepsilon_s^{i,h} - v_s^{i,h}$ in (51), and where $\lambda(\cdot)$ is a selection correction factor. As errors are normally distributed, $\lambda(\cdot)$ is simply (minus) the ratio of the standard normal density and cumulative distribution function, respectively (Heckman, 1979).

It thus follows that:

$$\mathbb{E}\left(\log H_s^{i,h}|P_s^{i,h}=1,\mathcal{X}_s^{i,h},Z_s^{i,h}\right)=e_{i,h}+f_{i,h}D_s+h'X_s^{i,h} \\ +g\mathbb{E}\left(\log\widetilde{W}_s^{i,h}|P_s^{i,h}=1,\mathcal{X}_s^{i,h},Z_s^{i,h}\right) \\ +g\nu\lambda\left(\text{Prob}\left(P_s^{i,h}=1|\mathcal{X}_s^{i,h}\right)\right).$$

To start with, suppose that the researcher knew

$$E_{it}=\mathbb{E}\left(\log\widetilde{W}_s^{i,h}|P_s^{i,h}=1,\mathcal{X}_s^{i,h},Z_s^{i,h}\right)$$

and $S_{it}=\lambda\left(\text{Prob}\left(P_s^{i,h}=1|\mathcal{X}_s^{i,h}\right)\right)$. Then g and h could be estimated by regressing log-hours on observed determinants of hours (D, X and individual dummies) and the two "regressors" E_{it} and S_{it}.

In practice, we replace the unknown quantities E_{it} and S_{it} by consistent estimates $\widehat{E}_{it}$ and $\widehat{S}_{it}$. The latter is easily computed, by replacing the probability of participation by the Probit estimate (as in Heckman, 1979). The former is more problematic.

A possibility is to adopt a semiparametric approach that places no restrictions on the conditional mean of observed log-wages. Thus, one could regress log-wages on a series in covariates and instruments whose degree depends on the sample size, and construct $\widehat{E}_{it}$ as the linear prediction in that regression (Newey, 1990).

We go half-way in that direction and regress, for labor market participants, the observed log-wage on the regressors $\mathcal{X}_s^{i,h}$, the instruments $Z_s^{i,h}$, and the selection correction factor $\widehat{S}_{it}$, and then construct $\widehat{E}_{it}$ as the linear prediction. Then, we regress the hours worked on labor market participants on $\mathcal{X}_s^{i,h}$, the selection correction factor, and the estimated conditional expectation $\widehat{E}_{it}$. This last regression delivers estimates $\widehat{g}$ and $\widehat{h}$.

Note that, in order for $\widehat{g}$ and $\widehat{h}$ to be consistent for g and h, respectively, we need that the model of the conditional expectation E_{it} be well-specified. Our linear specification including an inverse Mill's ratio can be viewed as a convenient approximation to the true conditional expectation.

Summary. To summarize, our estimation approach consists of three steps:

1. Regress participation using Probit. Predict the participation probabilities and compute the selection correction factor $\widehat{S}_{it}$.

2. Regress (observed) log-wages by OLS, including the job form instruments $Z_s^{i,h}$ and the predicted selection correction factor $\widehat{S}_{it}$

as additional controls. Predict the mean wage $\widehat{E}_{it}$, using a least-squares fit.

3. Regress log-hours, including $\widehat{E}_{it}$ and $\widehat{S}_{it}$ as controls.

Lastly, to account for the fact that E_{it} and S_{it} have been replaced by empirical estimates in step 3, we bootstrap the standard errors. We cluster the nonparametric bootstrap at the household level, hence computing clustered standard errors.[5,6]

6 Estimation results

We take the model to the Thai data. [**Details on the data; some information of regions**] Descriptive statistics on the sample may be found in Table 6.

Overall evidence. Table 1 reports the estimates of the coefficients of nonlabor income on participation and log-hours on the labor market. The results are separated into three categories. In the first category, year and month effects and regional interactions are included as controls in all equations. In the second category, interactions of time with village effects (which proxy for the Lagrange multipliers D_s) are also included. There are four regions and four villages by region in the dataset. Lastly, the third category adds individual fixed effects in order to control for unobserved individual heterogeneity.

For each category, the table shows the range of estimates across 18 specifications. Each specification is characterized by a sample selection criterion, a choice for the exclusion restriction, and the type of nonlabor income used. More precisely:

• There are three sample choices. In the first sample, we keep only individuals who worked at least 12 months on the labor market. In samples 2 and 3, we strengthen this restriction and keep individuals who worked at least 24 or 36 months on the labor market, respectively. This type of restriction is motivated by the need to precisely estimate individual fixed-effects.

[5] A similar approach can be used to estimate the model's coefficients when the outcomes of interest are participation and hours of work on the labor market or as a self-employed (i.e., on the farm). However, because the wages and instruments are only available for the participants to the wage labor market, we run the wage and hours regressions on the sample of wage workers only. Doing so is consistent with the economic model, which predicts that workers are indifferent between working on the wage labor market and working on the farm.

[6] In this version of the paper, we simply report the uncorrected clustered standard errors in the wage and hours equations.

• There are two types of exclusion restrictions: either lagged participation (consisting of 3 indicators: participation lagged 1 month, 2-6 months, and 7-12 months), or lagged participation of other household members (with the same 3 indicators). In this latter case we include lagged participation of the individual as an additional control in all equations.

• We use three types of nonlabor income: the first one does not include cash flows from production, the second one includes cash flows, and the third variable is a measure of "deficits", constructed using production data. [This needs to be clarified. Hero, could you help?]

Starting with the effects of household nonlabor income on participation, we see that all average marginal effects estimates reported in the top panel of Table 1 are negative. Moreover, as indicated in the columns labeled "< 0", the vast majority of the estimates are significantly so at the 2.5% level. Thus, a fall in household nonlabor income *increases* participation, even after controlling for aggregate effects. This is evidence of rejection of the risk-sharing hypothesis. However, the magnitude of the effects is small. Indeed, nonlabor income (which has been standardized to have unitary standard deviation to facilitate interpretation) has at most a 2% effect on the probability to participate to the labor market.

Turning to hours, we see in the bottom panel of Table 1 that the effects are also mostly negative. This is also evidence that the risk-sharing hypothesis is rejected. However, the magnitude of the effect is small in every specification, ranging from -3.6% to .08%. So, an increase in one standard deviation of nonlabor income implies at most a decrease in 3.6% in hours worked. In addition, the effects are significant in one third of the cases only, see the columns labelled "< 0".

To complement the results in Table 1, which give a summary of the effect of nonlabor income, we provide the complete regression results in Tables 7, 8 and 9. In particular, the results show that some parameters vary substantially across specifications. For example, this is the case of the wage elasticity in the hours equation. When including interactions of time and village dummies as controls, the elasticity estimate is 12%, marginally significant. However. when including in addition individual fixed effects the estimate drops to 3-4%, insignificant.[7] Note also that the effect of the selection correction factor in the log-hours equation (in Table 9) varies substantially between specifications.[8] In view of the

[7]Note that regressing log-hours on log-wages directly yields estimates of approximately -20%, highly significant. This is consistent with the presence of measurement error in hourly wages.

[8]The coefficient estimate is actually negative, suggesting somewhat counterintuitively that individuals who participate have lower potential hours than individuals

large variability of coefficient estimates across specifications, the stability of the effect of nonlabor income reported in Table 1 appears quite remarkable.

Effects by groups. Table 2 shows the effect of nonlabor income, for individuals who belong to "poor" and "non-poor" households, where poor households are defined as households whose total wealth over the period is below median. We see that poor households show more significant rejection of the risk sharing hypothesis. An increase in one standard deviation of nonlabor income decreases the probability to participate by about 2.5% for poor households, and by about 1% for non-poor households. Likewise, hours worked decrease by about 4% for poor households, versus 1% for the non-poor. Moreover, the difference between estimates for the poor and non-poor is mostly significant for participation, and significant in one third of the cases for hours, as shown by the rows labelled "difference" in the table.

The region-by-region estimates shown in Table 3 also show differences between the richer regions (7 and 49), and the poorer ones (27 and 53). In the latter, rejections of the risk sharing hypothesis are more striking, especially for log-hours in region 53 where the negative effect may exceed 10%.

Next, we study whether within-village networks have an effect on risk sharing. Table 4 shows the effect on nonlabor income on participation and hours, separately for individuals who belong to a network and individuals who are not connected. [**Details on how this information is constructed**] In the dataset, more than 70% of individuals belong to a network. The results show small differences between the two types of households, insignificant in all the specifications (see columns labelled "< 0"). Hence, insurance to income shocks seems to be driven by household wealth rather than connection to a network.

Lastly, in Table 5 we show the effects on nonlabor income on participation and hours devoted to any economic activity, and not only labor market activities. The results do not show evidence against risk sharing, as virtually all estimates reported in the table are insignificant from zero. As we mentioned before, however, there is a serious reverse causality issue in this case, so these estimates should be interpreted with caution.[9]

who do not participate.

[9] In particular, reverse causality should intuitively cause an upward bias in the estimates, possibly giving the impression that risk sharing holds while this is not the case.

References

[1] Borjas, G. (1980): "The Relationship Between Wages and Weekly Hours of Work: The Role of Division Bias," *Journal of Human Resources*, 15, 424-434.

[2] Heckman, J. (1979): "Sample Selection Bias as a Specification Error," *Econometrica*, 47, 73-92.

[3] Newey, W.K. (1990a): "Efficient Instrumental Variables Estimation of Nonlinear Models," *Econometrica*, 58, 809-837.

[4] J. Altonji, F. Hayashi, and L. Kotlikoff, "Is the Extended Family Altruistically Linked? Direct Tests Using Micro Data," American Economic Review, vol. 82(5), pages 1177-98, December. 1992

[5] S. Altug, and R. A. Miller. "Household Choices in Equilibrium," Econometrica, vol. 58(3), pages 543-70, May. 2000.

[6] P.A Chiappori, "Introducing Household Production in Collective Models of Labor Supply," Journal of Political Economy, vol. 105(1), pages 191-209, February, 1997

[7] P.A Chiappori and I. Ekeland, "Formal Models of Efficient Group Behavior The Economics and Mathematics of Aggregation", mimeo, 2008.

[8] P.A.Chiappori, K.Samphantharak, S. Schulhofer-Wohl, and R.M. Townsend. "Heterogeneity and Risk Sharing in Thai Villages", mimeo, 2005.

[9] P.A.Chiappori, R.M. Townsend, and H. Yamada. "Monte Carlo study on individual vs household consumption", mimeo, 2008.

[10] E. Duflo, and C. Udry, "Intrahousehold Resource Allocation in Cote d'Ivoire: Social Norms, Separate Accounts and Consumption Choices," NBER Working Papers 10498, National Bureau of Economic Research, 2004.

[11] F. Hayashi, J. Altonji, and L. Kotlikoff, "Risk-Sharing between and within Families," Econometrica, vol. 64(2), pages 261-94, March. 1996.

[12] M. Mazzocco, "Saving, Risk Sharing, and Preferences for Risk," American Economic Review, vol. 94(4), pages 1169-1182, September. 2004

[13] M. Mazzocco and S. Saini, "Testing Efficient Risk Sharing with Heterogeneous Risk Preferences: Semi-parametric Tests with an Application to Village Economies," mimeo, 2007.

[14] M.R. Rosenzweig, "Risk, Implicit Contracts and the Family in Rural Areas of Low-Income Countries." The Economic Journal 98 (393): 1148-1170 (December). 1988.

[15] R. Townsend, "Risk and Insurance in Village India," Econometrica, vol. 62(3), pages 539-91, May. 1994.

[16] R. Townsend, and H. Yamada. "The Summary of The Monthly Household Survey of Thailand on Labor Issues", mimeo, 2008.

Table 1: Test of the risk sharing hypothesis (labor market participation)

Time effects			+village interactions			+individual effects		
Range	Med.	< 0	Range	Med.	< 0	Range	Med.	< 0
Participation (average marginal effect)								
$-.020, -.006$	$-.012$	16/18	$-.020, -.007$	$-.011$	18/18	$-.021, -.011$	$-.016$	18/18
Log-hours (coefficient)								
$-.028, -.004$	$-.013$	3/18	$-.026, -.006$	$-.017$	6/18	$-.036, .008$	$-.012$	9/18

Note: Summary statistics of the coefficient of nonlabor income (standardized by its standard deviation) for 18 different specifications, see the text. "Time effects" means that yearly and monthly dummies interacted with regional dummies are included as controls, "+village interactions" means that the time dummies are interacted with village indicators, while "+individual effects" means that individual dummies are included in addition to time-village interactions. "Range" and "Med." refer to the range and median of the point-estimate across these 18 specifications, respectively, while "< 0" gives the proportion of specifications for which the effect is significantly negative at the 2.5% level. Note that, in this version of the paper, the standard errors in the hours equation do not reflect the fact that the selection correction factor has been estimated.

Table 2: Test of the risk sharing hypothesis, poor and non-poor households (labor market participation)

	Time effects			+village interactions			+individual effects		
	Range	Med.	< 0	Range	Med.	< 0	Range	Med.	< 0
				Participation (average marginal effect)					
Non-poor	−.012, −.002	−.007	11/18	−.014, −.002	−.007	10/18	−.015, −.007	−.011	13/18
Poor	−.030, −.015	−.024	18/18	−.031, −.017	−.024	18/18	−.035, −.019	−.026	18/18
Difference	−.021, −.012	−.017	16/18	−.021, −.011	−.017	14/18	−.022, −.008	−.015	10/18
				Log-hours (coefficient)					
Non-poor	−.023, .007	−.008	0/18	−.021, .003	−.011	0/18	−.019, .012	−.000	2/18
Poor	−.073, .002	−.037	8/18	−.078, −.002	−.038	8/18	−.082, −.001	−.042	12/18
Difference	−.076, .009	−.027	3/18	−.076, .006	−.029	6/18	−.070, −.005	−.043	11/18

Note: See note to Table 1. Here the coefficients reported are those of interactions of nonlabor income and dummies indicating whether the household has above or below median wealth ("non-poor" and "poor", respectively). The rows labeled "difference" report the range and median of the difference between the coefficients for "poor" and "non-poor" households, the column "< 0" indicating the proportion of specifications for which that difference is significantly negative at the 2.5% level.

Table 3: Test of the risk sharing hypothesis, by region (labor market participation)

	Time effects			+village interactions			+individual effects		
	Range	Med.	< 0	Range	Med.	< 0	Range	Med.	< 0
	Participation (average marginal effect)								
(7)	−.018, −.008	−.011	14/18	−.018, −.008	−.012	16/18	−.020, −.007	−.015	12/18
(27)	−.025, −.003	−.016	12/18	−.026, −.008	−.018	13/18	−.042, −.016	−.029	0/6
(49)	−.023, −.001	−.010	10/18	−.025, −.004	−.009	10/18	−.023, −.008	−.014	18/18
(53)	−.027, .015	−.001	4/18	−.025, .011	−.004	2/18	−.049, .002	−.023	9/18
	Log-hours (coefficient)								
(7)	−.031, .005	−.017	0/18	−.021, .006	−.011	0/18	−.041, .006	−.010	5/18
(27)	−.038, .012	−.018	2/18	−.049, .007	−.023	4/18	−.061, .012	−.016	4/18
(49)	−.022, .027	.009	0/18	−.012, .015	.001	0/18	−.028, .018	−.006	6/18
(53)	−.176, −.015	−.069	7/18	−.165, −.088	−.121	15/18	−.091, −.007	−.041	2/18

Note: See note to Table 1. Here the coefficients reported are those of interactions of nonlabor income and regional dummies.

Table 4: Test of the risk sharing hypothesis, households connected/not connected to a network (labor market participation)

	Time effects			+village interactions			+individual effects		
	Range	Med.	< 0	Range	Med.	< 0	Range	Med.	< 0
	Participation (average marginal effect)								
Network	$-.019, -.006$	$-.012$	15/18	$-.019, -.009$	$-.012$	18/18	$-.023, -.013$	$-.018$	18/18
No network	$-.019, -.005$	$-.009$	10/18	$-.020, -.004$	$-.010$	10/18	$-.019, -.005$	$-.009$	10/18
Difference	$-.001, .008$	.003	0/18	$-.001, .007$	.004	0/18	$-.005, .010$	.006	0/18
	Log-hours (coefficient)								
Network	$-.020, .008$	$-.009$	0/18	$-.029, .001$	$-.018$	4/18	$-.044, .008$	$-.015$	6/18
No network	$-.040, .010$	$-.028$	1/18	$-.029, .012$	$-.023$	0/18	$-.029, .007$	$-.008$	3/18
Difference	$-.035, .023$	$-.021$	0/18	$-.019, .033$	$-.005$	0/18	$-.004, .018$	.008	0/18

Note: See note to Table 1. Here the coefficients reported are those of interactions of nonlabor income and dummies indicating whether the household belongs to a network or not. The rows labeled "difference" report the range and median of the difference between the coefficients for households that do not belong to a network and households that do, the column "< 0" indicating the proportion of specifications for which that difference is significantly negative at the 2.5% level.

Table 5: Test of the risk sharing hypothesis, by region (participation to any economic activity)

	Time effects			+village interactions		
	Range	Med.	< 0	Range	Med.	< 0
	Participation (average marginal effect)					
(7)	$-.003, .007$	$-.002$	0/18	$-.003, .007$	$-.001$	0/18
(27)	$-.001, .006$	.002	0/18	$-.002, .005$	.002	0/18
(49)	$-.005, .007$	$-.000$	1/18	$-.004, .005$	.000	1/18
(53)	$-.009, .012$	.004	0/18	$-.005, .008$	.001	0/18
	Log-hours (coefficient)					
(7)	$.001, .013$	.008	0/18	$.005, .018$	.009	0/18
(27)	$-.042, .092$	.050	0/18	$-.043, .094$	.047	0/18
(49)	$-.023, .016$	$-.005$	0/18	$-.018, .004$	$-.006$	0/18
(53)	$-.032, .107$	.020	0/18	$-.069, .083$	$-.011$	2/18

Note: See note to Table 1. The dependent variables are participation and hours devoted to any economic activity. Here the coefficients reported are those of interactions of nonlabor income and regional dummies.

Table 6: Descriptive statistics [HERO: need to check]

	Demographics			
Region	(7)	(27)	(49)	(53)
Age (median)	45	43	44	46
Education = 4 or 5 (%) [HERO]	.21	.08	.11	.07
Gender (% male)	.44	.42	.46	.46
Household size (median)	4	4	4	4
	Participation, hours and income			
Region	(7)	(27)	(49)	(53)
Overall participation (%)	.72	.60	.76	.57
Participation (labor market, %)	.35	.24	.28	.14
Overall daily hours (median)	5.4	4.0	5.0	3.0
Daily hours (labor market, median)	6.1	4.3	5.9	3.4
Overall daily hours (labor market participants, median)	6.2	5.3	6.5	5.1
Hourly wage (median), in baht	158	60	101	57
Wealth (median), in million baht	1.19	.50	1.04	.43
	Job form data			
Region	(7)	(27)	(49)	(53)
Workers paid monthly versus daily (%)	.49	.37	.32	.74
Hourly wage (form, paid monthly, median)	172	91	121	67
Hourly wage (paid monthly, median)	206	100	138	86
Hourly wage (form, paid daily, median)	200	120	133	100
Hourly wage (paid daily, median)	138	42	108	23
	Nonlabor income			
Region	(7)	(27)	(49)	(53)
Nonlabor income (mean) [HERO]	148	23	193	48
Nonlabor income with cash flows (mean) [HERO]	203	69	242	79
"Deficits" (mean) [HERO]	-96	45	-92	13
	Counts			
Region	(7)	(27)	(49)	(53)
Number of individuals per village	180	161	162	170
Number of individuals	38,889	29,108	35,853	32,326
Individuals with participation $\geq$ 24 months	.77	.73	.85	.77
Individuals with labor market participation $\geq$ 24	.37	.24	.31	.11

Note: The nonlabor income data are truncated at the 1% and 99% percentiles.

Table 7: Labor market participation

Nonlabor income	−.058***(.018)	−.060***(.018)	−.097***(.019)	−.051***(.017)	−.052***(.017)	−.086***(.018)
Age (-20)	.265(.243)	.245(.240)	−.639*(.367)	.212(.256)	.177(.248)	−.659*(.368)
Age (21-40)	.312***(.074)	.344***(.079)	.063(.201)	.315***(.082)	.325***(.083)	.046(.201)
Age (41-60)	.298***(.072)	.326***(.081)	.311**(.174)	.287***(.079)	.305***(.088)	.369**(.161)
Household size	.035**(.015)	.034**(.015)	−.014(.024)	.035**(.016)	.033**(.016)	−.003(.030)
Gender	.129**(.061)	.129**(.063)	-	.132*(.077)	.127*(.081)	-
Education 2	.051(.082)	.067(.088)	-	.125(.096)	.157(.102)	-
Education 3	.297**(.138)	.298**(.145)	-	.392***(.141)	.410***(.150)	-
Education 4	.338***(.113)	.328**(.129)	-	.389***(.117)	.394***(.138)	-
Education 5	.489***(.173)	.477***(.174)	-	.541***(.183)	.551***(.183)	-
Lagged participation (1 month)	.862***(.053)	.859***(.053)	.886***(.057)	.812***(.063)	.813***(.064)	.856***(.071)
Lagged participation (2-6)	1.375***(.063)	1.401***(.062)	1.122***(.077)	1.537***(.071)	1.582***(.074)	1.287***(.087)
Lagged participation (7-12)	.617***(.052)	.626***(.040)	.477***(.059)	.594***(.051)	.587***(.050)	.463***(.056)
Household lagged partic. (1)	-	-	-	.757***(.075)	.746***(.076)	.719***(.099)
Household lagged partic. (2-6)	-	-	-	−.472***(.098)	−.481***(.098)	−.384***(.135)
Household lagged partic. (7-12)	-	-	-	−.111(.077)	−.103(.079)	.048(.152)
Month×region dummies	YES	YES	YES	YES	YES	YES
Month×village dummies	NO	YES	YES	NO	YES	YES
Individual dummies	NO	NO	YES	NO	NO	YES
Observations	30349	30146	23042	26282	26084	19366

Note: The dependent variable is labor market participation. The sample is conditional on working at least 2 years on the labor market. Standard errors clustered at the household level in parentheses. *,** ,***: *significant at* 10%, 5%, *and* 1%, *respectively.*

Table 8: Log-wages

Nonlabor income	−.006(.011)	−.009(.012)	−.005(.004)	−.003(.012)	−.004(.013)	−.003(.004)
Daily log-wage (job form)	.836***(.131)	.773***(.131)	.433***(.090)	.839***(.141)	.781***(.138)	.457***(.095)
Monthly log-wage (job form)	.539***(.078)	.498***(.078)	.258***(.054)	.542***(.083)	.504***(.082)	.277***(.057)
Age (-20)	.021(.125)	.027(.130)	.148(.138)	.045(.140)	.028(.145)	.110(.147)
Age (21-40)	−.010(.108)	.027(.116)	.010(.124)	.022(.124)	.026(.134)	−.033(.138)
Age (41-60)	.060(.112)	.104(.119)	.082(.111)	.083(.131)	.098(.138)	.032(.126)
Household size	−.014(.009)	−.010(.007)	.009(.010)	−.015*(.009)	−.014*(.008)	.011(.010)
Gender	−.124***(.034)	−.084**(.040)	-	−.128***(.039)	−.096**(.044)	-
Education 2	−.121(.078)	−.087(.075)	-	−.135(.090)	−.098(.093)	-
Education 3	−.128(.089)	−.067(.091)	-	−.138(.093)	−.076(.099)	-
Education 4	.018(.088)	.044(.096)	-	−.007(.101)	.025(.114)	-
Education 5	.229**(.103)	.246**(.108)	-	.196(.124)	.200(.130)	-
Lagged participation (1 month)	-	-	-	−.043(.055)	−.122*(.063)	−.031(.027)
Lagged participation (2-6)	-	-	-	−.130(.097)	−.267**(.115)	−.055(.048)
Lagged participation (7-12)	-	-	-	.081*(.044)	.028(.042)	.010(.029)
Selection factor	.070(.055)	.055(.042)	.015(.025)	−.008(.115)	−.238*(.144)	−.050(.044)
Month×region dummies	YES	YES	YES	YES	YES	YES
Month×village dummies	NO	YES	YES	NO	YES	YES
Individual dummies	NO	NO	YES	NO	NO	YES
Observations	18615	18461	11958	16558	16408	10314

Note: The dependent variable is market log-wages. The sample is conditional on working at least 2 years on the labor market. Standard errors clustered at the household level in parentheses. *,**,***: *significant at* 10%, 5%, *and* 1%, *respectively. Note that, in this version of the paper, standard errors do not reflect the fact that the selection correction factor has been estimated.*

Table 9: Log-hours

Nonlabor income	−.006(.011)	−.010(.010)	.008(.005)	−.016(.012)	−.020**(.010)	.003(.005)
Predicted log-wage	.085(.061)	.128*(.071)	.031(.174)	.079(.062)	.126*(.068)	.047(.170)
Age (-20)	.263**(.134)	.273**(.137)	.155(.203)	.329**(.132)	.380***(.135)	.137(.214)
Age (21-40)	.200*(.118)	.174(.115)	.109(.135)	.277**(.117)	.288**(.113)	.135(.146)
Age (41-60)	.121(.116)	.080(.109)	.046(.108)	.171(.117)	.164(.110)	.083(.118)
Household size	.024**(.012)	.022*(.013)	.007(.010)	.024**(.011)	.023*(.012)	.009(.011)
Gender	.096**(.041)	.063(.044)	-	.097**(.042)	.072(.044)	-
Education 2	−.005(.101)	−.047(.112)	-	−.002(.097)	−.040(.105)	-
Education 3	.114(.123)	.042(.136)	-	.105(.115)	.038(.125)	-
Education 4	.080(.109)	.028(.122)	-	.080(.101)	.016(.113)	-
Education 5	−.038(.116)	−.066(.132)	-	−.029(.107)	−.055(.123)	-
Lagged participation (1 month)	-	-	-	.061(.064)	.147**(.058)	.131***(.036)
Lagged participation (2-6)	-	-	-	.382***(.102)	.530***(.100)	.183***(.060)
Lagged participation (7-12)	-	-	-	.146***(.061)	.205***(.056)	−.008(.043)
Selection factor	−.749***(.078)	−.730***(.072)	−.410***(.041)	−.266*(.137)	−.034(.135)	−.137*(.080)
Month×region dummies	YES	YES	YES	YES	YES	YES
Month×village dummies	NO	YES	YES	NO	YES	YES
Individual dummies	NO	NO	YES	NO	NO	YES
Observations	18615	18461	11958	16558	16408	10314

Note: The dependent variable is market log-hours. The sample is conditional on working at least 2 years on the labor market. Standard errors clustered at the household level in parentheses. *,**,***: *significant at* 10%, 5%, *and* 1%, *respectively. Note that, in this version of the paper, standard errors do not reflect the fact that the selection correction factor and predicted log-wage have been estimated.*