Limits to Risk Sharing in Village Economies*

Sarolta Laczó[†]

August 2009

Abstract

This paper examines how well models of risk sharing explain the allocation of consumption in a community. The models of perfect risk sharing, autarky, and risk sharing with limited commitment are estimated in a structural manner. We extend the approach of Ligon, Thomas, and Worrall (2002) in several ways. First, we allow households' preferences to depend on observable household characteristics. Second, this paper allows for unobservable individual effects. Third, measurement error is dealt with explicitly in the structural model by using a simulated maximum likelihood estimator. Finally, Vuong's (1989) tests are applied to statistically compare the models. We provide evidence that both heterogeneity in preferences and limitations to the enforcement of informal insurance contracts are important in explaining the consumption allocation. Simulations of the effects of redistributive policies are conducted based on the structural estimations, using household survey data from rural Pakistan.

Keywords: risk sharing, limited commitment, preference heterogeneity, structural microeconometrics

JEL codes: C52, D10, D52

*Very preliminary, comments and suggestions welcome. This paper has greatly benefitted from discussions with Richard Blundell, Pierre Dubois, Luigi Guiso, Thierry Magnac, David Pacini, Alessandro Tarozzi, and Judit Vall Castello. All errors are mine. I thank the International Food Policy Research Institute (IFPRI) for making the data available. Generous funding from Agence Nationale de la Recherche (ANR-06-BLAN-0104) is gratefully acknowledged.

†Toulouse School of Economics (Gremaq), Manufacture des Tabacs, Aile Jean-Jacques Laffont, MF007, 21 allée de Brienne, 31000 Toulouse, France, sarolta.laczo@univ-tlse1.fr.

1 Introduction

Households living in rural areas of low-income countries often face a great amount of risk. Revenue from agricultural production is usually low and volatile, as a result of extreme weather conditions, like erratic monsoon rains in Pakistan. Further, outside job opportunities are often lacking. In addition, financial instruments to insure against consumption fluctuations are scarcely available. In such an environment households in a community have to rely on one another for insurance.

There exists ample empirical evidence that households in poor villages do not fully share the risks they face, while they do achieve a remarkable amount of insurance without formal contracts (Townsend, 1994, Grimard, 1997, Dubois, 2000, Dercon and Krishnan, 2003a,b, and others). The observed partial insurance may be due to private information or lack of commitment, because informal insurance contracts are not enforceable by third parties. The latter is arguably the better assumption, since households in small communities in developing countries can observe shocks faced by their neighbors (bad crop, or illness), but no authority exists to enforce insurance contracts. The model of risk sharing with limited commitment has been developed by Thomas and Worrall (1988), Coate and Ravallion (1993), Kocherlakota (1996), and Ligon, Thomas, and Worrall (2002), and its implications are supported by mounting empirical evidence (see Fafchamps, 1999, Ligon, Thomas, and Worrall, 2002, 2002, Attanasio and Ríos-Rull, 2000, Foster and Rosenzweig, 2001, and others). In this paper, informal insurance transfers are required to be voluntary, or self-enforcing, but we assume away informational problems.

This paper compares the model of risk sharing with limited commitment, with and without preference heterogeneity, to the benchmarks of perfect risk sharing and autarky. We aim to statistically evaluate how well these models are able to explain the distribution of consumption, taking income as exogenous. Since the estimation is done in a structural manner, the effects of policies on the consumption allocation can be simulated. Thus this research may provide guidance for the evaluation and design of redistributive policies or micro-insurance programs, taking into account existing informal arrangements to share risk.

This paper is most related to the work of Ligon, Thomas, and Worrall (2002), LTW hereafter, and extends their empirical approach in several ways. First, we allow for preference heterogeneity across households. In particular, households' utility may depend on observable covariates. We are then able to discuss how relative risk aversion depends on household characteristics. Second, this paper allows for unobserved individual effects. Third, measurement error is dealt with explicitly in the structural model. Finally, we perform statistical tests to compare the models, both with and without preference heterogeneity. In particular, we apply likelihood ratio-based tests introduced by Vuong (1989). The tests' great advantage is that we do not have to assume that any model is correctly specified.

The extension to introduce preference heterogeneity in the case of perfect risk sharing has been explored by several papers. Dubois (2000) specifies an isoelastic utility function, and allows the coefficient of relative risk aversion to depend on observables, as well as for multiplicative preference shocks. Schulhofer-Wohl (2007) uses data on risk aversion, and finds evidence that occupational choice is affected by risk preferences. He argues that this should be taken into account when evaluating how well people are able to mitigate the adverse effects of risk they face. He then constructs a new test of perfect risk sharing, where he captures heterogeneity in risk preferences by a nuisance parameter. Mazzocco and Saini (2008) construct nonparametric tests of perfect risk sharing allowing for preference heterogeneity. The present paper contributes to this strand of literature by looking at the limited commitment case as well, but considers only parametric models.

This paper also contributes to the literature on explaining consumption inequality given income inequality. Krueger and Perri (2006) show that, as a result of partial insurance, observed cross-sectional consumption inequality is smaller than income inequality. Partial insurance is modeled allowing for limited commitment, as here, but the model is then calibrated rather than estimated. The authors argue that consumption inequality increased less than income inequality in the United States over the last few decades because more income risk induces more informal insurance. On the other hand, Blundell, Pistaferri, and Preston (2008) document that income shocks have become less persistent, and thereby easier to insure against. The present paper only allows for transitory shocks, but builds and estimates a structural model of how consumption is allocated, given income.

This paper first details the different theoretical models of risk sharing. Households are assumed to be infinitely lived and risk averse. They face some exogenous risk on their income each period. The distribution from which incomes are drawn is common knowledge ex ante, as well as income realizations ex post. The only way households may mitigate the adverse effects of risk they face is to insure one another. We examine insurance across states of the world, and not time, thus savings are assumed away. Perfect risk sharing means that all idiosyncratic risks are insured, and aggregate risk is borne efficiently. In other words, households pool their income, then distribute it according to predetermined Pareto-weights, and less risk averse households bear more of the uninsurable aggregate risk (Borch, 1962, Wilson, 1968). However, formal insurance contracts are often not available in rural villages in developing countries, and the perfect risk sharing solution might not be self-enforcing. That is, a household with a high income realization today may decide not to contribute, but to quit the risk sharing arrangement instead, and stay in autarky thereafter.

The risk sharing with limited commitment model characterizes the case where enforcement constraints may be binding. This is arguably a good description of informal insurance arrangements among households in rural communities, or members of a family (Mazzocco, 2007).[1] This paper talks in details about the model, and discusses how to solve it using numerical dynamic programming. A co-state variable, namely, the relative weight of households in the social planner's objective, has to be introduces to rewrite the problem in a recursive form, following the ideas of Marcet and Marimon (1998). The solution of the model is fully characterized by a set of state-dependent intervals on the relative Pareto-weights, or, the ratio of marginal utilities (LTW). Given preferences and the distribution from which incomes are drawn, the optimal interval for each income state can be solved for numerically. Then, the consumption allocation predicted by the model is computed trying to keep the ratio of

[1]The model has also been used in a wide variety of other economic contexts, including risk sharing between an employee and an employer (Thomas and Worrall, 1988), and countries (Kehoe and Perri, 2002). Further, Schechter (2007) adopts the same model to examine the interaction between a farmer and a thief, and Dixit, Grossman, and Gul (2000) use a similar model to examine cooperation between opposing political parties.

marginal utilities as close as possible to the one from the previous period, while respecting the interval for the income state realized today.

Then, in section 3, the empirical models are set up. In particular, we are interested in how the consumption allocation today is explained by past consumption, incomes today and in the past, and household characteristics, according to the different models of risk sharing. More precisely, we model the changes in each household's consumption relative to mean consumption in the community. Maximum likelihood estimators are derived assuming that the multiplicative errors in the measurement of consumption are log-normally distributed. The estimation of the perfect risk sharing model, as well as the benchmark of autarky, are relatively straightforward. The main difficulty in the limited commitment case is that the updating of the co-state variable depends on unobservable individual effects and measurement error. We show that the predicted consumption allocation is not affected by the individual effects. To deal with measurement error, a simulated estimator is used. In addition, the solution of the model can only be approximated. Therefore, we compare the model of risk sharing with limited commitment with a given precision to other models of risk sharing.[2] Vuong's (1989) tests are appropriate in this context, see Fernández-Villaverde, Rubio-Ramírez, and Santos (2006).

The data comes from an income-consumption survey conducted by the International Food Policy Research Institute (IFPRI) in rural Pakistan. Almost 1000 households have been interviewed over 12 rounds in 46 villages in 4 districts of Pakistan. We use data from 6 villages in Faisalabad district in Punjab. The survey contains detailed information on household characteristics, consumption of food and other nondurable goods, and income from different sources, including crop production, wage labor, small businesses, and remittances from abroad. For the purposes of this paper, we only need a measure of consumption, income, and some household characteristics.

We estimate and compare five models using these data: perfect risk sharing with and without preference heterogeneity, autarky, and the model of risk sharing with limited commitment

[2] Robustness checks are preformed comparing models with different levels of precision in approximating the solution.

with and without preference heterogeneity. We do not assume that any of the models are correctly specified, and use both the nested and non-tested tests proposed by Vuong (1989) to compare the models. Both preference heterogeneity and limits to the enforcement of risk sharing contracts turn out to be important. We then simulate the effects of a redistributive policy, taking into account existing informal insurance arrangements.

The rest of the paper is structured as follows. First, the theoretical models are discussed. Then, section 3 details the empirical models. Section 4 talks about the data used. Section 5 contains the estimation results for the structural models, both with and without preference heterogeneity, as well the statistical tests to compare the models. Robustness checks are also performed. Section 6 conducts policy simulations. In section 7, some extensions to the empirical models are discussed. Concluding remarks are presented in section 8.

2 Models of risk sharing

Suppose that there are N infinitely-lived, risk-averse households in a community. They consume a private and perishable consumption good c. Each household i maximizes her expected lifetime utility,

$$E_0 \sum_t \delta^t u_i(c_{it}),$$

where E_0 is the expected value at time 0 calculated with respect to the probability measure describing the common beliefs, $\delta \in (0,1)$ is the (common) discount factor, and c_{it} is the consumption of household i at time t. The instantaneous preferences of household i are described by the utility function

$$u_i(c_{it}) = \exp(\eta_i) \frac{c_{it}^{1-\sigma_i} - 1}{1-\sigma_i}, \tag{1}$$

where $\sigma_i > 0$ is the coefficient of relative risk aversion of household i, and η_i accounts for preference heterogeneity unrelated to risk aversion.

Suppose that random income, denoted Y_i for household i, is independently and identically distributed (i.i.d.) over time for each household, and is drawn from some discrete distribution F_{Y_i} for household i. Let s_t denote the state of the world that describes the income realizations of all households in the community at time t. The distribution F_{Y_i}, $\forall i$, is common knowledge ex ante, as well as income realization ex post at each time t. That is, there are no informational problems. Note also that income is exogenous. In other words, the effect of risk on choices among different income generating processes is ignored. In addition, we study consumption smoothing across states of the world, not across time periods, thus savings are assumed away.

This section considers three models in turn. First, it talks about the model of perfect risk sharing. Second, subsection 2.2 mentions the benchmark of autarky. Third, the model of risk sharing with limited commitment is detailed in subsection 2.3.

2.1 Perfect risk sharing

To find the first best, the Pareto-optimal allocations, we solve the social planner's problem. The (utilitarian) social planner maximizes a weighted sum of households' excepted lifetime utilities,

$$\max_{\{c_{it}(s_t)\}} \sum_i \lambda_i \sum_{t=1}^{\infty} \sum_{s_t} \delta^t \pi\left(s_t\right) u_i\left(c_{it}\left(s_t\right)\right),$$

where λ_i is the (initial) weight of household i in the social planner's objective, δ is the discount factor, and $\pi\left(s_t\right)$ is the probability of state s_t occurring; subject to the resource constraint

$$\sum_i c_{it}\left(s_t\right) = \sum_i y_{it}\left(s_t\right), \forall s_t, \forall t,$$

where $y_{it}\left(s_t\right)$ is the income of household i at time t and state s_t.

The well-known result that

$$\frac{u_k'\left(c_{kt}\left(s_t\right)\right)}{u_i'\left(c_{it}\left(s_t\right)\right)} = \frac{\lambda_i}{\lambda_k}, \forall s_t, \forall t, \tag{2}$$

that is, the ratio of marginal utilities for any two households i and k is constant over time and across states of the world, follows from the first order conditions of the social planner's problem (Borch, 1962, Wilson, 1968). (2) also means that all idiosyncratic risks are insured away, and households share aggregate risk efficiently. With the utility function (1), the first order condition (2), for any s_t and t, is

$$\frac{\exp\left(\eta_k\right) c_{kt}^{-\sigma_k}}{\exp\left(\eta_i\right) c_{it}^{-\sigma_i}} = \frac{\lambda_i}{\lambda_k}. \tag{3}$$

2.2 Autarky

When households stay in autarky, the problem is trivial, since savings have been assumed away. The model predicts that

$$c_{it}\left(s_t\right) = y_{it}\left(s_t\right), \forall s_t, \forall i, \forall t. \tag{4}$$

Let $U_i^{aut}\left(s_t\right)$ denote the expected lifetime utility, or, the value function, of household i in autarky at state s_t and time t. It can be computed by solving the Bellman equation

$$U_i^{aut}\left(s_t\right) = u_i\left(y_{it}\left(s_t\right)\right) + \delta \sum_{s_{t+1}} \pi\left(s_{t+1}\right) U_i^{aut}\left(s_{t+1}\right). \tag{5}$$

In the present i.i.d. case, we may also write the value function as

$$U_i^{aut}\left(s_t\right) = u_i\left(y_{it}\left(s_t\right)\right) + \frac{\delta}{1-\delta} \sum_{s} \pi\left(s\right) u_i\left(y_i\left(s\right)\right).$$

2.3 Risk sharing with limited commitment

We may write the problem as follows. The social planner maximizes a weighted sum of households' expected lifetime utilities,

$$\max_{\{c_{it}(s^t)\}} \sum_i \lambda_i \sum_{t=1}^{\infty} \sum_{s^t} \delta^t \pi\left(s^t\right) u_i\left(c_{it}\left(s^t\right)\right),$$

where $\pi\left(s^t\right)$ is the probability of history $s^t = \left(s_1, s_2, ..., s_t\right)$ occurring, and $c_{it}\left(s^t\right)$ denotes the consumption of household i when history s^t has occurred; subject to the resource constraints

$$\sum_i c_{it}\left(s^t\right) \leq \sum_i y_{it}\left(s_t\right), \forall s^t, \forall t, \tag{6}$$

and the enforcement constraints,

$$\sum_{r=t}^{\infty} \sum_{s^r} \delta^{r-t} \pi\left(s^r \mid s^t\right) u_i\left(c_{ir}\left(s^r\right)\right) \geq U_i^{aut}\left(s_t\right), \forall s^t, \forall t, \forall i, \tag{7}$$

where $\pi\left(s^r \mid s^t\right)$ is the probability of history s^r occurring given that history s^t has occurred up to time t, and the right hand side has been defined in equation (5). Note that even if income is i.i.d., consumption may depend on the whole history of income realizations.

Denoting the multiplier on the enforcement constraint of household i (7) by $\delta^t \pi(s^t) \mu_i(s^t)$, and the multiplier on the resource constraint (6) by $\delta^t \pi(s^t) \rho(s^t)$ when history s^t has occurred, the Lagrangian is

$$\sum_{t=1}^{\infty} \sum_{s^t} \delta^t \pi\left(s^t\right) \left[\sum_i \lambda_i u_i\left(c_{it}\left(s^t\right)\right) \right.$$
$$+ \mu_i\left(s^t\right) \left(\sum_{r=t}^{\infty} \sum_{s^r} \delta^{r-t} \pi\left(s^r \mid s^t\right) u_i\left(c_{ir}\left(s^r\right)\right) - U_i^{aut}\left(s_t\right) \right)$$
$$\left. + \rho\left(s^t\right) \left(\sum_i y_{it}\left(s_t\right) - c_{it}\left(s^t\right) \right) \right].$$

Using the ideas of Marcet and Marimon (1998), the Lagrangian can also be written in the form

$$\sum_{t=1}^{\infty} \sum_{s^t} \delta^t \pi\left(s^t\right) \left[\sum_i M_i\left(s^{t-1}\right) u_i\left(c_{it}\left(s^t\right)\right) \right.$$
$$\left. + \mu_i\left(s^t\right) \left(u_i\left(c_{it}\left(s^t\right)\right) - U_i^{aut}\left(s_t\right) \right) + \rho\left(s^t\right) \left(\sum_i y_{it}\left(s_t\right) - c_{it}\left(s^t\right) \right) \right],$$

where $M_i(s^t) = M_i(s^{t-1}) + \mu_i(s^t)$ with $M_i(s^0) = \lambda_i$ (see also Kehoe and Perri, 2002). In words, $M_i(s^t)$ is the initial weight of household i plus the sum of the Lagrange multipliers on her enforcement constraints along the history s^t.

The first order condition with respect to $c_{it}(s^t)$ is

$$\delta^t \pi\left(s^t\right) M_i\left(s^t\right) u_i'\left(c_{it}\left(s^t\right)\right) - \rho\left(s^t\right) = 0. \tag{8}$$

There are also standard first order conditions relating to the resource and enforcement constraints, with complementarity slackness conditions. Let us consider two households sharing risk, household i and k. Household k can be thought of as the rest of the community, as in LTW, or a typical household. Combining the first order conditions (8) for these two households for history s^t at time t, we have

$$\frac{u_k'\left(c_{kt}\left(s^t\right)\right)}{u_i'\left(c_{it}\left(s^t\right)\right)} = \frac{M_i\left(s^t\right)}{M_k\left(s^t\right)} = \frac{\lambda_i + \mu_i\left(s^1\right) + \mu_i\left(s^2\right) + \ldots + \mu_i\left(s^t\right)}{\lambda_k + \mu_k\left(s^1\right) + \mu_k\left(s^2\right) + \ldots + \mu_k\left(s^t\right)} \equiv x_i\left(s^t\right), \tag{9}$$

where $x_i(s^t)$ can be thought of as the relative Pareto-weight assigned to household i when history s^t has occurred.

The vector of relative weights $x(s^t)$, with elements $x_i(s^t)$ defined in (9), can be used as an additional co-state variable in order to rewrite the problem in a recursive form (Marcet and Marimon, 1998). The current income state s_t does not tell us everything we need to know about the past, only (s_t, x_{t-1}) does, where x_{t-1} is the relative weight, equal to the ratio of marginal utilities, inherited from the previous period. Denote by x_t the new relative weight to be found at time t. The solution consists of policy functions for the consumption allocation and the new relative weight, with support over the extended state space (s_t, x_{t-1}). That is, $c_{it}(s_t, x_{t-1})$, $\forall i$, and $x_t(s_t, x_{t-1})$ are to be determined. At last, the value functions can be defined recursively as

$$V_i(s_t, x_{t-1}) = u_i(c_{it}(s_t, x_{t-1})) + \delta \sum_{s_{t+1}} \pi(s_{t+1}) V_i(s_{t+1}, x_t(s_t, x_{t-1})). \tag{10}$$

The solution can be fully characterized by a set of state-dependent intervals on the relative weight, or, the ratio of marginal utilities, x, that give the possible relative weights in each income state (LTW). Denote the interval for state s by $[\underline{x}^s, \overline{x}^s]$. Suppose that last period the ratio of marginal utilities was x_{t-1}, and today the income state is s. Today's ratio of marginal utilities, x_t, is determined by the following updating rule (LTW):

$$x_t = \begin{cases} \overline{x}^s & \text{if} \quad x_{t-1} > \overline{x}^s \\ x_{t-1} & \text{if} \quad x_{t-1} \in [\underline{x}^s, \overline{x}^s] \\ \underline{x}^s & \text{if} \quad x_{t-1} < \underline{x}^s \end{cases} \tag{11}$$

Numerical dynamic programming allows us to solve for the optimal intervals, and thereby the consumption allocation, given the income processes, utility functions, and discount rates of the two households. After a sufficient number of periods, the initial relative weight in the social planner's objective only matters if perfect risk sharing is self-enforcing (Kocherlakota, 1996).

Remember that F_{Y_i} (F_{Y_k}) denotes the distribution from which household i's (k's) income is drawn. Then, we can solve numerically for x_{it} given,

$$(y_{it}, y_{kt}, x_{i,t-1}; \eta_i, \eta_k, \sigma_i, \sigma_k, \delta, F_{Y_i}, F_{Y_k}).$$

Details are in the Appendix. Once we know x_{it}, the first order conditions (9) and the resource constraint (6) give the consumption allocation predicted by the model.

3 Empirical models

Let us first specify the elements of the utility function (1). η_i can be thought of as an unobserved individual effect. As for σ_i, assume that it is a linear function of observables, in particular,

$$\sigma_i = 1 + z_i'\beta,$$

where β is a parameter vector to be estimated, and z_i represents a vector of time-invariant observable characteristics of household i. Note that z_i does not contain an (additional) constant, as in Dubois (2000). A normalization is needed, because consumption risk borne by each household is determined by her risk tolerance relative to average risk tolerance in the community. Further, if the coefficient on the constant is a free parameter, then, taking all households as risk neutral, any consumption allocation would be Pareto optimal.[3] Remember that the above theoretical models assume perfect information, thus the preferences of each household are known to everybody, but the econometrician only observes z_i, $\forall i$.

Assume that consumption is measured with a multiplicative measurement error that is log-normally distributed. Let c_{it}^* denote consumption observed by the econometrician, and let $\exp(\varepsilon_{it})$ be the multiplicative measurement error in household i's consumption at time t. Then, we may write

$$c_{it}^* = \exp(\varepsilon_{it})\, c_{it},$$

where ε_{it} is independently and identically distributed (i.i.d.) across households and time, and $\varepsilon_{it} \sim N(0, \gamma^2)$, where γ^2 is to be estimated. Note that true consumption c_{it} is observed by all households in the community. Measurement error in income is ignored for now, and is introduced as an extension in section 7.

We model the allocation of observed consumption, $c_t^* \equiv (c_{1t}^*, ..., c_{it}^*, ..., c_{Nt}^*)$, for $t = 2, ..., T$, determined by the history of income realizations, time-constant household characteristics, observed consumption at time 1, c_1^*, and parameters. In mathematical terms, we are interested in how the following conditional density could be specified based on the above models of risk

[3] This is because marginal utility is always 1 for a risk-neutral households, thus any consumption allocation would keep the ratio of marginal utilities constant.

sharing:

$$f\left(c_T^*, ..., c_2^* \mid c_1^*, y_T, ..., y_1, Z; \beta, \delta, \gamma^2, F_Y, \eta, \lambda\right), \tag{12}$$

where y_t, for $t = 1, ..., T$, is the vector of income realizations for households at time t, $Z = [z_1, ..., z_i, ..., z_N]'$ is the matrix of household observables for all households, $\theta = (\beta, \delta, \gamma^2, F_Y)$ are the structural parameters to be estimated,[4] where $F_Y = F_{Y_1,...,Y_i,...,Y_N}$ is the joint distribution of households' incomes, and the vectors η and λ are nuisance parameters.

This paper considers models that explain the allocation of consumption in each community at each time t, but not how aggregate consumption changes over time. Therefore, we model each household's consumption relative to mean consumption in the community.

To deal with the unobservable, time-constant parameters, namely, the individual effects η and the initial Pareto-weights in the social planner's objective λ, we first difference. This means trying to explain how the consumption allocation changes from one period to the next. In the limited commitment case, our estimator is then similar to the 'changes-in-shares' estimator of LTW, that was found to be the best at capturing how consumption reacts to income.

For the limited commitment case, LTW have shown that the updating rule (11) holds for both the 2- and N-household case. In the empirical part, they approximate the N-household economy by looking at each household i sharing risk with the 'rest of the community.' We follow their approach in this paper. This results in important gains in computation time. We often call the rest of the community household k. Household k can also be thought of as the chief of the community, coordinating transfers.

An additional issue is how to specify the preferences of household k. Let us subtract for each household from each observable in the utility function its community mean. This then means that the preferences of a typical household in the community are described by an isoelastic utility function, with coefficient of relative risk aversion equal to 1. That is, $u_k(c_{kt}) = \log c_{kt}$. Normalize also the Pareto-weight of household k to 1, that is, $\lambda_k = 1$. This is without loss of generality, since only relative Pareto weights matter. Further, we

[4]Below θ often denotes a subset of these parameters, and is used as a short form for 'structural parameters to be estimated.'

assume that c_{kt} is well measured, since the variance of the measurement error in mean consumption in the community is only a fraction of the variance of the measurement error in each household's consumption. This assumption is only for notational simplicity. Think of explanatory variables in the utility function as deviations from their community mean hereafter, abusing notation.[5]

The next three subsections detail in turn how the model of perfect risk sharing (subsection 3.1), autarky (3.2), and risk sharing with limited commitment (3.3) are estimated. The estimations are done using (simulated) maximum likelihood estimators, and Vuong's (1989) tests are applied to statistically compare the models. Subsection 3.4 says more on model selection.

3.1 Perfect risk sharing

In the case of perfect risk sharing, the current consumption allocation should only depend on current and not past exogenous variables. It depends neither on the discount factor, nor on the distribution from which incomes are drawn. However, it depends on the time-constant unobservables, η and λ. Thus (12) can be written as

$$\prod_{t=2,\dots,T} f\left(c_t^* \mid y_t, Z; \beta, \gamma^2, \eta, \lambda\right). \tag{13}$$

Further, c_t^* only depends on today's income realizations through aggregate income.

To fix ideas, let us consider household i and the average household, k. Taking the logarithm of the first order condition with respect to (true) consumption for these two households, equation (3), noting that $\sigma_k = 1$ and $\lambda_k = 1$, we get

$$\sigma_i \log c_{it} - \eta_i - \log c_{kt} = \log \lambda_i.$$

Replacing for σ_i and rearranging gives

$$\log\left(\frac{c_{it}}{c_{kt}}\right) = -z_i'\beta \log c_{it} + \eta_i + \log \lambda_i. \tag{14}$$

[5] Note also that, when preferences are homogeneous, meaning $\beta = 0$ here, the coefficient of relative risk aversion is normalized to 1, that is, $u_i(c_{it}) = \log c_{it}$, $\forall i$.

In terms of measured consumption c_{it}^*, (14) reads

$$\log\left(\frac{c_{it}^*}{c_{kt}}\right) = -z_i'\beta \log c_{it}^* + \eta_i + \log \lambda_i + (1 + z_i'\beta)\,\varepsilon_{it}.$$

Now, let us take first differences to eliminate $\eta_i + \log \lambda_i$. Doing so and rearranging yields

$$\log\left(\frac{c_{it}^*}{c_{kt}}\right) = \log\left(\frac{c_{i,t-1}^*}{c_{k,t-1}}\right) - z_i'\beta \log\left(\frac{c_{it}^*}{c_{i,t-1}^*}\right) + (1 + z_i'\beta)\,(\varepsilon_{it} - \varepsilon_{i,t-1})\,. \tag{15}$$

Let $\psi^2(\theta) \equiv 2\,(1 + z_i'\beta)^2\,\gamma^2$, and

$$d_{it}^{prs}(\theta) \equiv \left[\log\left(\frac{c_{it}^*}{c_{kt}}\right) - \log\left(\frac{c_{i,t-1}^*}{c_{k,t-1}}\right) + z_i'\beta \log\left(\frac{c_{it}^*}{c_{i,t-1}^*}\right)\right] / \psi(\theta)\,.$$

Then, we may write the likelihood of observation it as

$$L_{it}^{prs}(\theta) = \phi\left(d_{it}^{prs}(\theta)\right),$$

where ϕ is the density of the standard normal distribution. Finally, the (pseudo) maximum likelihood estimator (MLE) maximizes

$$\ell^{prs}(\theta) = \sum_{i=1}^{N}\sum_{t=2}^{T} \log \phi\left(d_{it}^{prs}(\theta)\right), \tag{16}$$

with respect to θ, that is, the vector β and the variance γ^2. The model is also estimated without preference heterogeneity for comparison. This means setting $\beta = 0$, thus the only parameter that remains to be estimated is γ^2.

We do not assume that the model is correctly specified and compute the variance-covariance matrix of the estimated parameters without assuming that the information matrix equality holds. We also take into account serial correlation. In particular, the variance-covariance matrix is estimated by $\hat{A}^{-1}\hat{B}\hat{A}^{-1}$, where

$$\hat{A} = \textstyle\sum_{i=1}^{N}\sum_{t=2}^{T} -\nabla_\theta^2 \ell_{it}(\hat{\theta}) \quad \text{and} \quad \hat{B} = \sum_{i=1}^{N}\sum_{t=2}^{T} \hat{s}_{it}\hat{s}_{it}' + \sum_{i=1}^{N}\sum_{t=2}^{T}\sum_{r\neq t} \hat{s}_{ir}\hat{s}_{it}',$$

where $\hat{s}_{it} = \nabla_\theta \ell_{it}(\hat{\theta})'$ is the score evaluated at the estimated parameters, and where the second term in the expression for $\hat{B}$ accounts for serial correlation (Wooldridge, 2002). Both the first and second derivatives of the log-likelihood function can be computed analytically here.

3.2 Autarky

Taking the logarithm of (4), first differencing, and introducing measured consumption gives

$$\log\left(\frac{c^*_{it}}{c^*_{i,t-1}}\right) = \log\left(\frac{y_{it}}{y_{i,t-1}}\right) + \left(\varepsilon_{it} - \varepsilon_{i,t-1}\right).$$

The consumption of household i relative to mean consumption is

$$\log\left(\frac{c^*_{it}}{c_{kt}}\right) = \log\left(\frac{c^*_{i,t-1}}{c_{k,t-1}}\right) + \log\left(\frac{y_{it}}{y_{kt}}\right) - \log\left(\frac{y_{i,t-1}}{y_{k,t-1}}\right) + \left(\varepsilon_{it} - \varepsilon_{i,t-1}\right), \tag{17}$$

where we have just added and subtracted $\log c_{kt} = \log y_{kt}$ and $\log c_{k,t-1} = \log y_{k,t-1}$ to have the same dependent variable in the equation to be estimated as above. In terms of the allocation of consumption within a community, the autarky model says that the change in the consumption share of household i should be the same as the change in her income share.

Let

$$d^{aut}_{it}(\theta) = \left[\log\left(\frac{c^*_{it}}{c^*_{i,t-1}}\right) - \log\left(\frac{y_{it}}{y_{i,t-1}}\right)\right] / \sqrt{2\gamma^2}.$$

The likelihood of observation it is

$$L^{aut}_{it}(\theta) = \phi\left(d^{aut}_{it}(\theta)\right),$$

and the log-likehood function to be maximized is

$$\ell^{aut}(\theta) = \sum_{i=1}^{N}\sum_{t=2}^{T} \log\phi\left(d^{aut}_{it}(\theta)\right). \tag{18}$$

The only parameter to be estimated is γ^2. We allow for misspecification and serial correlation when computing the variance of the estimated parameter.

3.3 Risk sharing with limited commitment

In the limited commitment case, the (true) ratio of marginal utilities from last period, x_{t-1}, the co-state variable in the recursive version of the model, is a sufficient statistic for everything that happened in the past, including the initial condition. In other words, instead of conditioning on the whole history of income realizations y^t, and the initial Pareto-weights in the social planner's objective λ, it is sufficient to condition on current income realizations y_t

and x_{t-1}. However, unlike in the perfect risk sharing case, the consumption allocation may also depend on the discount factor δ, and the distribution from which incomes are drawn F_Y. Thus (12) becomes

$$\prod_{t=2,\ldots,T} f\left(c_t^* \mid y_t, Z; \beta, \delta, \gamma^2, F_Y, \eta, x_{t-1}\right), \tag{19}$$

where x_{t-1} has elements $x_{i,t-1}$, which can be thought of as the true relative weight of household i with respect to household k at time $t-1$. Note that, along with η, x_{t-1} is not observed. Therefore, knowing the density (19) as specified by the model of risk sharing with limited commitment is not enough, we have to determine the distribution conditioning on x_{t-1}^* with elements

$$x_{i,t-1}^* = \frac{\left(c_{i,t-1}^*\right)^{1+z_i'\beta}}{c_{k,t-1}},$$

the observable ratio of marginal utilities at time $t-1$, instead of x_{t-1} with elements

$$x_{i,t-1} = \frac{\left(c_{i,t-1}\right)^{1+z_i'\beta}}{\exp\left(\eta_i\right) c_{k,t-1}}.$$

Since x_{t-1}^* can be easily computed given observables and parameters, we may write

$$\prod_{t=2,\ldots,T} f\left(c_t^* \mid c_{t-1}^*, y_t, Z; \beta, \delta, \gamma^2, F_Y, \eta\right).$$

Note, however, that with measurement error, all past values of consumption could be informative on $x_{i,t-1}$. For tractability, we only deal with the above density.

The first order condition in the limited commitment case can be written as

$$\log\left(\frac{c_{it}}{c_{kt}}\right) = -z_i'\beta \log c_{it} + \eta_i + \log x_{it}, \tag{20}$$

replacing x_{it} for λ_i in (14). According to the model of risk sharing with limited commitment, given $x_{i,t-1}$, preferences, current income realizations, and the distribution from which incomes are drawn, we can solve numerically for x_{it}, $\forall i$, that we denoted $x_t\left(s_t, x_{t-1}\right)$, see (10) and (11). Let the function $g()$ denote this relationship, that is, $x_{it} = g\left(y_t, x_{i,t-1} \mid Z; \theta, \eta_i\right)$, with $\theta = \left(\beta, \delta, F_Y\right)$. Replacing for x_{it} in (20) gives

$$\log\left(\frac{c_{it}}{c_{kt}}\right) = -z_i'\beta \log c_{it} + \eta_i + \log g\left(y_t, x_{i,t-1} \mid Z; \theta, \eta_i\right), \tag{21}$$

Remember that, instead of c_{it}, the econometrician observes $c^*_{it} = \exp(\varepsilon_{it})\, c_{it}$. Then, in terms of observable consumption (21) is

$$\log\left(\frac{c^*_{it}}{c_{kt}}\right) = -z_i'\beta \log c^*_{it} + \eta_i + \log g\left(y_t, x_{i,t-1} \mid Z; \theta, \eta_i\right) + \left(1 + z_i'\beta\right) \varepsilon_{it}. \tag{22}$$

3.3.1 The likelihood

The main econometric issue is that unobservables, namely, individual effects and measurement error influence the updating of the state variable. That is, among the arguments of $g()$ in equation (22), η_i is not observed, and instead of $x_{i,t-1}$ only

$$x^*_{i,t-1} = \exp(\eta_i) \left(\exp\left(\varepsilon_{i,t-1}\right)\right)^{1+z_i'\beta} x_{i,t-1}$$

is observed.

Let us first deal with the individual effects η, and assume that we know the realization of the measurement error in household i's consumption at time $t-1$, denoted $\varepsilon^j_{i,t-1}$, drawn from the distribution of $\varepsilon_{i,t-1}$, $N\left(0, \gamma^2\right)$. To deal with the individual effects, we show that

$$\begin{aligned} g\left(y_t, x_{i,t-1} \mid Z, \varepsilon^j_{i,t-1}; \theta, \eta_i\right) &= \exp(-\eta_i)\, g\left(y_t, x^*_{i,t-1} \mid Z, \varepsilon^j_{i,t-1}; \theta\right) \\ &= \exp(-\eta_i) \frac{\hat{c}_{it}\left(y_t, c^*_{i,t-1}, Z, \varepsilon^j_{i,t-1}; \theta\right)^{1+z_i'\beta}}{\hat{c}_{kt}\left(y_t, c^*_{i,t-1}, Z, \varepsilon^j_{i,t-1}; \theta\right)}, \end{aligned}$$

where $\hat{c}_{it}\left(y_t, c^*_{i,t-1}, Z, \varepsilon^j_{i,t-1}; \theta\right)$ and $\hat{c}_{kt}\left(y_t, c^*_{i,t-1}, Z, \varepsilon^j_{i,t-1}; \theta\right)$ is the consumption of household i and k, respectively, predicted by the model, normalizing $\eta_i = 0$. Replacing this in equation (22), η_i drops out. The fact that the function $g()$ is homogeneous of order one in η_i is the direct consequence of the following proposition.

Proposition 1. $\hat{c}_{it}\left(y_t, c^*_{i,t-1}, Z, \varepsilon^j_{i,t-1}; \theta\right) = \hat{c}_{it}\left(y_t, c^*_{i,t-1}, Z, \varepsilon^j_{i,t-1}; \theta, \eta_i\right)$. *That is, the consumption allocation predicted by the model of risk sharing with limited commitment does not depend on the individual effects.*

Proof. To see this, let us take a closer look at the enforcement constraint of some household

i, that can be written as (7) in general. Replacing the utility function (1) in (7) gives

$$\exp(\eta_i) \frac{c_{it}(s^t)^{1-\sigma_i} - 1}{1-\sigma_i} + \sum_{r=t+1}^{\infty} \sum_{s^r} \delta^{r-t} \pi(s^r) \exp(\eta_i) \frac{c_{ir}(s^r)^{1-\sigma_i} - 1}{1-\sigma_i} \geq$$
$$\geq \exp(\eta_i) \frac{y_{it}(s_t)^{1-\sigma_i} - 1}{1-\sigma_i} + \sum_{r=t+1}^{\infty} \sum_{s_r} \delta^{r-t} \pi(s_r) \exp(\eta_i) \frac{y_{ir}(s_r)^{1-\sigma_i} - 1}{1-\sigma_i}. \tag{23}$$

Both sides can be divided by $\exp(\eta_i)$, thereby eliminating the individual effects. When no enforcement constraint is binding, we are back to perfect risk sharing, where $\exp(\eta_i)$ appears multiplicatively on both sides of $x_{it} = x_{i,t-1}$. □

Using (5) and (10), and replacing for the utility function with $\sigma_i = 1 + z_i'\beta$, the enforcement constraint of household i at time t, that the predicted consumption allocation has to satisfy, can be written in a recursive form as

$$\frac{c_{it}(s_t, x_{t-1})^{-z_i'\beta} - 1}{-z_i'\beta} + \delta \sum_{s_{t+1}} \pi(s_{t+1}) V_i(s_{t+1}, x_t(s_t, x_{t-1})) \geq$$
$$\geq \frac{y_i(s_t)^{-z_i'\beta} - 1}{-z_i'\beta} + \delta \sum_{s_{t+1}} \pi(s_{t+1}) V_i^{aut}(s_{t+1}). \tag{24}$$

This inequality is to be used in the numerical solution of the model, with

$$c_{it}(s_t, x_{t-1}) = \hat{c}_{it}\left(y_t, c^*_{i,t-1}, Z, \varepsilon^j_{i,t-1}; \theta\right) \quad \text{and} \quad x_t(s_t, x_{t-1}) = g\left(y_t, x^*_{i,t-1} \mid Z, \varepsilon^j_{i,t-1}; \theta\right).$$

In the perfect risk sharing case, the predicted consumption allocation was independent of δ, the discount factor. The question now is whether we can identify this parameter in the case of risk sharing with limited commitment. Proposition 2 states that the answer is yes, if some but not perfect insurance occurs. Denote by $\overline{\delta}$ the discount factor such that, $\forall \delta \geq \overline{\delta}$, perfect risk sharing is self-enforcing, and denote by $\underline{\delta}$ the discount factor such that, $\forall \delta \leq \underline{\delta}$, all households stay in autarky.[6]

Proposition 2. *The parameter δ is identified if $\delta \in \left(\underline{\delta}, \overline{\delta}\right)$, that is, if some informal insurance is achieved and at least one enforcement constraint binds.*

[6]LTW have shown that $\overline{\delta}$ and $\underline{\delta}$ exist.

Proof. Let us prove this for the case with homogeneous risk preferences. The argument for the heterogeneous case is similar. Compared to the perfect risk sharing case, additional information can only come from binding enforcement constraints. Suppose that at time t households i's enforcement constraint is binding.[7] Let us rewrite (24) with equality and with $z_i'\beta = 1 - \sigma = 0$. Simple algebra then gives

$$\log y_{it}(s_t) - \log c_{it}(s_t, x_{t-1}) = \delta \sum_{s_{t+1}} \pi(s_{t+1}) \left[V_i(s_{t+1}, x_t(s_t, x_{t-1})) - V_i^{aut}(s_{t+1}) \right],$$

where the left hand side is the utility cost of the transfer household i makes today, and the right hand side is welfare gain of sharing risk according to the informal insurance contract rather than staying in autarky in the future. If the right hand side is strictly monotonic and continuous in δ, and only this constraint ever binds, we could perfectly match household i's consumption at time t from the data, with a unique appropriately chosen δ.

The expected future gain of insurance is strictly increasing in δ, for $\delta \in (\underline{\delta}, \overline{\delta})$, since a higher δ relaxes all enforcement constraints. Note that as δ approaches 1, perfect risk sharing, the first best, is self-enforcing by the well-known folk theorem. On the other extreme, when it is close to 0, no voluntary transfers are made. In between, the higher δ is, the closer transfers get to their first-best level. In other words, when δ is higher, more informal insurance is achieved, and consumption is smoother across income states. In other words, a higher δ means a better enforcement technology.

It is easy to see that $V_i^{aut}(s_{t+1})$ is continuous in δ. As for $V_i(s_{t+1}, x_t)$, LTW have shown that the limits of the optimal state-dependent intervals, that fully characterize the solution of the model, are continuous in δ (see LTW, page 219). Since $V_i(s_{t+1}, x_t)$ is a continuous function of these limits, it is itself continuous in δ. It follows that one binding enforcement constraint identifies δ. □

Let us now consider measurement error. Remember that $\varepsilon_{i,t-1}^j$ denotes the realization of the measurement error in household i's consumption at time $t-1$. To deal with the fact that measurement error enters the updating of the state variable, we first write the likelihood

[7] We also implicitly assume that observed consumption is lower than income for household i at time t.

of each observation conditional on $\varepsilon^j_{i,t-1}$. Then, averaging the conditional likelihood over J draws, we integrate $\varepsilon_{i,t-1}$ out. That is, in the case of risk sharing with limited commitment with measurement error, we use a simulated (pseudo) maximum likelihood estimator (SMLE).

Conditional on $\varepsilon^j_{i,t-1}$, (22) becomes

$$\log\left(\frac{c^*_{it}}{c_{kt}}\right) = \log\left(\frac{\hat{c}_{it}\left(y_t, c^*_{i,t-1}, Z, \varepsilon^j_{i,t-1};\theta\right)}{\hat{c}_{kt}\left(y_t, c^*_{i,t-1}, Z, \varepsilon^j_{i,t-1};\theta\right)}\right) - z_i'\beta \log\left(\frac{c^*_{it}}{\hat{c}_{it}\left(y_t, c^*_{i,t-1}, Z, \varepsilon^j_{i,t-1};\theta\right)}\right) + \left(1+z_i'\beta\right)\varepsilon^j_{it}. \tag{25}$$

Note that, even though we have not taken first differences explicitly, when perfect risk sharing is self-enforcing, (25) is equivalent to (15). This is because

$$\log\left(\frac{\hat{c}_{it}\left(y_t, c^*_{i,t-1}, Z, \varepsilon^j_{i,t-1};\theta\right)}{\hat{c}_{kt}\left(y_t, c^*_{i,t-1}, Z, \varepsilon^j_{i,t-1};\theta\right)}\right) + z_i'\beta \log \hat{c}_{it}\left(y_t, c^*_{i,t-1}, Z, \varepsilon^j_{i,t-1};\theta\right)$$
$$= \log\left(\frac{c^*_{i,t-1}}{c_{k,t-1}}\right) + z_i'\beta \log c^*_{i,t-1} - \left(1+z_i'\beta\right)\varepsilon_{it-1}$$

in that case. Similarly, we get back the estimating equation of autarky, equation (17), if $\delta \leq \underline{\delta}$.

Let $\psi^2(\theta) \equiv \left(1+z_i'\beta\right)^2\gamma^2$, and

$$d^{lc}_{it}\left(y_t, c^*_{i,t-1}, Z, \varepsilon^j_{i,t-1};\theta\right) = \left[\log\left(\frac{c^*_{it}}{c_{kt}}\right) - \log\left(\frac{\hat{c}_{it}\left(y_t, c^*_{i,t-1}, Z, \varepsilon^j_{i,t-1};\theta\right)}{\hat{c}_{kt}\left(y_t, c^*_{i,t-1}, Z, \varepsilon^j_{i,t-1};\theta\right)}\right) + z_i'\beta \log\left(\frac{c^*_{it}}{\hat{c}_{it}\left(y_t, c^*_{i,t-1}, Z, \varepsilon^j_{i,t-1};\theta\right)}\right) - \left(1+z_i'\beta\right)\varepsilon^j_{i,t-1}\right]/\psi(\theta). \tag{26}$$

Then, the likelihood of observation it given $\varepsilon^j_{i,t-1}$ is

$$L^{lc}_{it}(\theta) = \phi\left(d^{lc}_{it}\left(y_t, c^*_{i,t-1}, Z, \varepsilon^j_{i,t-1};\theta\right)\right).$$

Making J draws for $\varepsilon^j_{i,t-1}$, the simulated likelihood of observation it is

$$L^{lc}_{it}(\theta) = \frac{1}{J}\sum_{j=1}^{J}\phi\left(d^{lc}_{it}\left(y_t, c^*_{i,t-1}, Z, \varepsilon^j_{i,t-1};\theta\right)\right).$$

Finally, the simulated log-likelihood function to be maximized is

$$\ell^{lc}(\theta) = \sum_{i=1}^{N}\sum_{t=2}^{T}\log\left(\frac{1}{J}\sum_{j=1}^{J}\phi\left(d^{lc}_{it}\left(y_t, c^*_{i,t-1}, Z, \varepsilon^j_{i,t-1};\theta\right)\right)\right). \tag{27}$$

Allowing for misspecification, the SMLE consistently estimates the pseudo-true values of the parameters and is asymptotically normal, if both the number of *it* observations, that we denote M, and the number of simulations J tend to infinity, and $\sqrt{M}/J \to 0$ (see Gouriéroux and Monfort, 1997, for example). When computing the variance-covariance matrix, the information matrix equality is not assumed to hold, and serial correlation is taken into account. The score and the hessian are computed numerically.

3.3.2 Estimation

The estimation is done in three steps. A preliminary step (i) involves estimating the distribution from which income is drawn, F_{Y_i}, for each household i. Then, (ii) the inside optimization computes the consumption allocation predicted by the model, given observable covariates and parameters. Finally, (iii) the log-likelihood (27) is maximized over the remaining structural parameters, $\theta = (\beta, \delta, \gamma^2)$. Now we turn to the details of each of these steps.

(i) The discrete distribution from which income is drawn, F_{Y_i}, $\forall i$, has to be estimated. This cannot be done in general, because the time dimension of the panel is not large enough (maximum 12 for each household). We assume that all households face the same multiplicative risk, and estimate the common distribution of the multiplicative risk nonparametrically. In particular, we create quantiles over observed incomes divided by individual mean income. Mean income can be thought of as a proxy for the income generating capacity of the household. For the rest of the community, quantiles are computed over mean income in the community. We allow for 7 income states for each household, and 4 for the rest of the community. The income states for household i inputed to the model for household i is the estimated quantiles over the multiplicative risk times mean income of household i.

(ii) We have to solve the inside optimization to find the consumption allocation predicted by the model. The Bellman equation (10) is solved by iteration. A grid is defined over the continuous state variable x_i. At iteration h, we solve for the new consumption values in states where an enforcement constraint is binding using (24) with equality, while the ratio of marginal utilities stays constant in other states. The values from iteration $h - 1$ are kept for $V_i(s_{t+1}, x_t)$ in (24). At the first iteration the values of perfect risk sharing can be used. We keep iterating until the policy function converges, that is, the optimal state-dependent intervals on x_i do not change. The algorithm to solve for the constrained-efficient risk sharing contract, given observables and structural parameters, does not impose much additional difficulty relative to the case without preference heterogeneity and heterogeneity in Y_i, except for computation time. Computation time is proportional to the number of households. The appendix gives more details on the algorithm. This step leads to the predicted consumption values, $\hat{c}_{it}\left(y_t, c^*_{i,t-1}, Z, \varepsilon^j_{i,t-1}; \theta\right)$ and $\hat{c}_{kt}\left(y_t, c^*_{i,t-1}, Z, \varepsilon^j_{i,t-1}; \theta\right)$, $\forall i$, to be replaced in (25).

(iii) The maximum-likelihood estimates of the structural parameters, the vector β, δ, and γ^2, are obtained by iteration between the dynamic program that solves for the predicted consumption allocation and the likelihood maximization routine. For comparison, the model is also estimated without preference heterogeneity. There only δ and γ^2 are to be estimated. The likelihood maximization is done, after a crude manual grid search, by a standard optimization algorithm available in R, namely, BFGS with bounds (L-BFGS-B),[8] which is a quasi-Newton method. The preliminary grid search is necessary, since we can only identify δ on the interval $\left(\underline{\delta}, \overline{\delta}\right)$, see Proposition 2.

The estimation of the risk sharing with limited commitment model involves both simulation and approximation. We take the number of simulations $J = 100$, which is a reasonable choice, given that we will have close to 1500 *it* observations, thus $J \gg \sqrt{M}$. Computation time is only moderately increased when adding an additional draw, because the optimal intervals, that fully characterize the solution of the model, do not have to be recomputed. Below we check that our results are robust with respect to changing J.

The continuous co-state variable x_i has to be discretized, and we use a 30-point grid. Computation time is approximatively proportional to the number of income states, that we take $7 \times 4 = 28$, and the square of the number of gridpoints on x_i. Increasing the number of gridpoints would be beneficial in better approximating the true solution of the model. We

[8]See www.r-project.org.

are limited by the cost in terms of computation time. Additional approximation error may come from the fact that we limit the number of iterations when solving the model.

Using Vuong's tests below, we allow for approximation errors, as a form of misspecification. We may say that what is compared is the limited commitment model with 30 gridpoints on x_i (instead of a continuous x_i). In a recent paper, Ackerberg, Geweke, and Hahn (2008) argue that, in terms of asymptotic properties of the maximum likelihood estimator, approximation error in dynamic computed models have similar effects as a limited number of simulations. Fernández-Villaverde, Rubio-Ramírez, and Santos (2006) have recently proposed a way to test whether approximation errors are important when estimating computed dynamic models. They propose to compare models with different precisions, e.g. in terms of the number of gridpoints, using Vuong's test. They provide Monte Carlo evidence that the method indeed works, assuming that the model estimated is the correct one. In this paper we do not assume that any of the models are correctly specified, and it is possible that in the case of risk sharing with limited commitment, a higher number of gridpoints would lead to a decrease in the model's fit to data. As a robustness check, below we look at whether our results are affected when changing the number of gridpoints.

3.4 Model selection

To statistically compare the above models, we use model selection tests introduced by Vuong (1989). Vuong proposes likelihood ratio-based statistics to compare nested and non-nested models. These statistics allow us to test the null hypothesis that two competing models are equally close to the true data generating process, against the alternative that one model is closer. Neither model has to be correctly specified. Further, we can correct for the number of parameters to be estimated.

The tests are based on the difference between the log likelihood values of the two models being compared. Suppose that we want to compare model 1 and model 2 using M observations. Denote the log likelihood of observation m for model 1 (2) at the estimated parameter vector $\hat{\theta}^1$ ($\hat{\theta}^2$) by ℓ_m^1 (ℓ_m^2). The adjusted likelihood ratio is defined as

$$LR = \sum_{m=1}^{M} \left(\ell_m^1 - \ell_m^2\right) - K,$$

where K is a correction factor that can account for the number of parameters, for example. Denote the number of parameters to be estimated by p (q) for model 1 (2). We take $K = p - q$, that corresponds to the Akaike (1973) information criterion.

If the two models are non-nested, then, under the null hypothesis that the two models are equally close to the true model,

$$\frac{LR}{\sqrt{M}\hat{\omega}} \Rightarrow N\left(0,1\right),$$

where $\hat{\omega}$ is the estimated standard deviation of the likelihood ratio, that is,

$$\hat{\omega}^2 = \frac{1}{M}\sum_{m=1}^{M}\left(\ell_m^1 - \ell_m^2\right)^2 - \left(\frac{1}{M}\sum_{m=1}^{M}\left(\ell_m^1 - \ell_m^2\right)\right)^2,$$

and where $\Rightarrow$ means convergence in distribution. If the two models are nested, and we want to allow for the possibility that the unconstrained model is not correctly specified, then under the null

$$2LR \Rightarrow M_{p+q}\left(\cdot;\hat{\kappa}\right),$$

where $M_{p+q}\left(\cdot;\hat{\kappa}\right)$ is the cumulative distribution function of a weighted sum of $p+q$ χ^2 distributions with degrees of freedom equal to 1 (see Vuong, 1989, page 313).[9] The p-values of the weighted χ^2 distribution are simulated. We do 100,000 replications.

We compare five models: perfect risk sharing with heterogeneous preferences (PRS$^{\text{het}}$), perfect risk sharing with homogeneous preferences (PRS$^{\text{hom}}$), autarky (AUT), risk sharing with limited commitment with heterogeneous preferences (LC$^{\text{het}}$), and risk sharing with limited commitment with homogeneous preferences (LC$^{\text{hom}}$). LC$^{\text{het}}$ nests all the other models.

[9] The weights $\hat{\kappa}$ can be computed by finding the real, nonzero eigenvalues of the matrix

$$\begin{bmatrix} -\hat{B}^1(\hat{A}^1)^{-1} & -\hat{B}^{1,2}(\hat{A}^2)^{-1} \\ \hat{B}^{2,1}(\hat{A}^1)^{-1} & \hat{B}^2(\hat{A}^2)^{-1} \end{bmatrix},$$

where $\hat{A}^1 = \sum_{i=1}^{N}\sum_{t=2}^{T} -\nabla_\theta^2 \ell_{it}^1$, $\hat{B}^1 = \sum_{i=1}^{N}\sum_{t=2}^{T} \hat{s}_{it}^1 \hat{s}_{it}^{1\prime}$, similarly for model 2, and $\hat{B}^{1,2} = \hat{B}^{2,1\prime} = \sum_{i=1}^{N}\sum_{t=2}^{T} \hat{s}_{it}^1 \hat{s}_{it}^{2\prime}$.

LC^{hom} nests PRS^{hom} and AUT, but LC^{hom} and PRS^{het} are non-nested. PRS^{het} nests PRS^{hom}. PRS^{het} and AUT, as well as PRS^{hom} and AUT are non-nested.

A caveat to this procedure is that we have to assume that observations are i.i.d. This can be considered as a form of misspecification. Alternatively, assuming that consumption *growth* is measured with error, as in Cochrane (1991), for example, the errors of the first differenced equations are i.i.d. In this case, however, we couldn't take measurement error properly into account in the limited commitment case, since we would have to draw $\varepsilon^j_{i,t-1}$ from a random walk. Rivers and Vuong (2002) propose tests for the non-i.i.d. case, but only for non-nested models, unfortunately, while here the model of risk sharing with limited commitment with heterogeneous preferences nests all other models.

4 Data

The data comes from an income-consumption survey conducted by the International Food Policy Research Institute (IFPRI) in rural Pakistan between July 1986 and September 1989. Almost 1000 households were interviewed over 12 rounds in 46 villages in 4 districts of Pakistan. The districts were not chosen randomly: 3 are the least-developed districts in their respective provinces (Attock in Punjab, Badin in Sind, and Dir in North-West Frontier Province), while the 4^{th} is a more prosperous district (Faisalabad in Punjab). Then, in each district, two markets were chosen, and villages were randomly selected from a stratified sample based on distance from these markets. Finally, households were chosen randomly within each village. Attrition seems to be due to administrative problems, and not households' self-selection, and we assume that attrition is random. In each household, both the male and female heads were interviewed. In addition, village questionnaires were also administered, that give information about prices, for example. For further details see Alderman and Garcia (1993). Due to computation time constraints, we only present results for Faisalabad district in Punjab. Pakistani Punjab has well developed factor and product markets, compared to the poor semi-arid areas on which much work on risk sharing in developing countries is based (Kurosaki and Fafchamps, 2002). Therefore, finding limits to risk sharing in Punjab makes it likely that similar constraints to insurance exist elsewhere.

This dataset is attractive for the purposes of the present paper for several reasons. First of all, both the cross-sectional and the time dimension is relatively big compared to other similar datasets. Second, we can examine both small and larger communities (villages and districts). Further, consumption data were collected from the female head of the household, while income data from the male head. Thus the assumption that measurement error in consumption is independent of the income measure is more compelling than usual.

4.1 Variables used

For the purposes of this paper, we need measures of consumption, income, and some household characteristics. To measure consumption we use both food consumption and nondurable consumption per adult equivalent. Nondurable consumption is constructed as the sum of food consumption, expenditures on clothing, hygiene items, tobacco, and cinema. To compute the adult equivalent household size we use the same age-gender weights as Townsend (1994).[10] Consumption is weekly and is expressed in 1986 Pakistani rupees. In 1986 about 16 rupees were worth 1 US dollar, that is about 2 2009 US dollars.

Income is constructed as the sum of net income from crop production, net income from poultry and livestock, net income from craft work, net income from produce from orchards, income from assets (hiring out bullocks, tractor, thresher, land, income from mills owned, and selling water), wage income minus wages paid for hired labor (that is not used in agricultural production), transfers from the state, and transfers from abroad, that is, all transfers from outside of the community (transfers from within Pakistan from friends, relatives, or religious organizations, typically the local mosque, are thus excluded). Medical expenditures and education investment are subtracted (as in Gourinchas and Parker, 2002). Income is then divided by the adult equivalent size of the household. Income is expressed in 1986 Pakistani rupees per week, as consumption.

[10] These weights are: 1 for adult males, 0.9 for adult females, 0.94 and 0.83 for males and females aged 13-18, respectively, 0.67 for children aged 7-12, 0.52 for children aged 4-6, 0.32 for children aged 1-3, and 0.05 for infants below 1 year of age.

We consider three variables when specifying preference heterogeneity, namely, age of the head of the household at time 1 (`age`), gender, and education. Household heads are almost exclusively male in the sample, thus we construct a measure of the gender composition of households. In particular, `gender` equals the time-average of the proportion of women among adults. To measure education, a categorical variable is used based on final schooling achievement of the head of the household. In particular, `education` equals 1 if the head is illiterate, 2 if he has gone to primary school or learnt to read, 3 or 4 if he has attended middle or secondary school (including technical studies), respectively, and 5 if he went to college or university. Education serves as a proxy for household wealth. We do not include this later variable, because many wealth items are not measured at time 1, thus endogeneity concerns arise. The correlation between `education` and adult equivalent wealth is 0.40.[11]

We delete observation *it* if consumption is missing, income is missing, or any income component is outside some reasonable range. We delete households whose head is above 80 years of age, or if the head changes over the three-year period of the interviews. Further, we delete villages in which we have observations on less than 5 households. Thereafter we are left with a sample size of 168 households and 1850 observations for Faisalabad district. Table 1 presents descriptive statistics.

[11]Wealth is the sum of the value of land owned, houses, other assets (like TV sets, watches), tools, and livestock. These are measured at different dates, unfortunately mostly not at time 1, so we take an average, lacking a better alternative. We value all items using prices available in the survey, except for land. We value land starting from the median rent per acres for different types of land (rainfed, or canal irrigated, for example). Then we use data from Renkow (1993) that rents are 2% and 2.6% of land prices for rainfed and better-quality land, respectively, in the 1986-1989 period in Punjab.

Table 1: Descriptive statistics, Faisalabad district

Variable	Mean	Sd	Min	Max	Observations
Food consumption[a]	342.45	177.01	46.07	1399.5	1850
Aeq. food consumption	56.86	29.25	8.775	356.04	1850
Log(aeq. food consumption)	3.925	0.4829	2.172	5.875	1850
Nondurable consumption[a]	781.42	836.88	54.15	9839.4	1850
Aeq. nondurable consumption	126.36	126.86	10.59	1430.0	1850
Log(aeq. nondurable consumption)	4.539	0.7455	2.360	7.265	1850
Income[a]	162.00	555.01	−1089.0	6168.7	1850
Aeq. income	25.50	86.86	−162.12	934.66	1850
Household size	9.042	4.444	1	41	1850
Aeq. household size	6.417	2.376	1	16.89	1850
Age of head[b]	50.51	13.85	18	80	168
Gender[c]	0.4815	0.1202	0	0.8333	168
Education[d]	1.839	1.205	1	5	168

[a]Measured in 1986 Pakistani rupees per week. 16 rupees = 1 US dollar in 1986.
[b]As of the time of the first interview.
[c]Proportion of women among adults.
[d]Categorical variable describing schooling achievement of the head of the household.

On average, daily adult equivalent nondurable consumption is just over 1 1986 US dollar, which is about 2 2009 US dollars. Households are poor even in the more prosperous Faisalabad district. Measured income is only a fraction of measured consumption, which reflects general underreporting, and the fact that agricultural production was hard hit by bad weather conditions during the years of the survey. Households are big, they have 9 members on average. Often the extended family forms one economic unit. Almost 60% of household heads are illiterate, and only about 14% have schooling achievement above middle school.

For the structural estimations, we delete the 5% extreme consumption and income observations in the pooled panel. Finally, aggregate income should be equal to aggregate consumption in the community, since savings have been assumed away. To achieve this, income is rescaled so that aggregate income be equal to aggregate consumption at each t.

4.2 Existing evidence

The present dataset has been used by a number of papers to examine risk sharing. Dubois (2000) constructs a nondirectional test of perfect risk sharing based on overidentifying restrictions implied by the model. Allowing for preference heterogeneity, he is able to reject that households share risk perfectly. He also shows that sharecropping contracts are used to achieve better insurance. Dubois, Jullien, and Magnac (2008) also provide evidence that perfect risk sharing is not achieved, and both short-term formal and informal insurance contracts are important.

On the other hand, Ogaki and Zhang (2001) are not able to reject perfect risk sharing for the vast majority of villages examined, when relative risk aversion is not constant. They use both the ICRISAT Indian data and the dataset from Pakistan used in this paper. The authors argue that earlier tests of perfect risk sharing do not take into account the possibility of decreasing relative risk aversion. In other words, the coefficient of relative risk aversion may depend on wealth. The present paper allows risk aversion to depend on more observables, and compares the perfect risk sharing model to a well-specified alternative, namely the model of risk sharing with limited commitment.

5 Structural estimation and model selection results

We consider five models, risk sharing with limited commitment with and without preference heterogeneity (LC^{het} and LC^{hom}, respectively), perfect risk sharing with and without heterogeneous preferences (PRS^{het} and PRS^{hom}, respectively), and autarky (AUT); and for each model, we first take communities to be districts, then villages. We expect commitment problems to be less severe within villages, since a smaller, close community is expected to better enforce informal contracts. Note, however, that modeling household consumption relative to the village average is less ambitious, since we do not try to explain how consumption is allocated between villages within a district. In both cases, Vuong's tests are performed to statistically compare the five models.

Let us specify

$$z_i'\beta = \beta_1\,\texttt{education}_i + \beta_2\,\texttt{age}_i + \beta_3\,\texttt{gender}_i, \tag{28}$$

thus the coefficient of relative risk aversion is allowed to depend on three variables. First of all, note that if the coefficient of relative risk aversion is higher, that means that the marginal utility of consumption is lower, given that consumption is greater than 1. It also means that the household bears less of the aggregate risk faced by the community.

We now discuss what is the expected sign of the parameters in (28). Most empirical studies find that women are more risk averse, see Jianakoplos and Bernasek (1998) for example. Exceptions include the seminal work of Binswanger (1980), who finds no difference in risk aversion between men and women. Note that we only observe household consumption here, thus the proportion of women only affects household risk aversion if women's preferences are different and they influence household decisions. Evidence on the effect of age and education on risk aversion is mixed.[12] Thinking of education as a proxy for wealth, we may discuss the sign of β_1 based on how risk aversion is expected to change with wealth. There is agreement in the literature on the fact that *absolute* risk aversion decreases with wealth. However, the effect of wealth on *relative* risk aversion is debated. Both Arrow (1965) and Pratt (1964) hypothesized that relative risk aversion increases with wealth, while empirical evidence on the matter is mixed (see Halek and Eisenhauer, 2001, and references therein).

We first estimate the models taking aggregate consumption in the district as constant. Then, aggregate consumption in the village is assumed to be exogenous, and we look at each household's consumption relative to the village mean. A district is a rather small geographical area, and we may suppose that people are directly or indirectly connected within a district. This assumption would be much stronger considering households in different districts. However, informational problems should hinder risk sharing more in the district than within villages. We do not allow all coefficients to be different by village, to keep sample size and statistical power higher. The computations have been done using the software R, see www.r-project.org.

[12] Guiso and Paiella (2008) find that risk aversion is independent of wealth and deceases with education, based on the willingness to pay for a hypothetical risky security. Wang and Hanna (1997) document that investment in risky assets increases with age, controlling for other individual characteristics, which implies that elder people are less risk averse. On the other hand, Palsson (1996) finds that relative risk aversion increases with age using data on portfolio decision of Swedish households. Shaw (1996) finds that better educated individuals take more risk. Based on insurance data, Halek and Eisenhauer (2001) identify a positive relationship between education and risk aversion. In the laboratory experiment of Holt and Laury (2002) on choosing between risky prospects, risk aversion is independent of both age and education.

5.1 Main results

Table 2 shows the structural estimation results for nondurable consumption for all the models considering households sharing risk within Faisalabad district. Table 3 then looks at risk sharing within villages.

Table 2: Risk sharing in Faisalabad district, adult-equivalent *nondurable* consumption

Model	LC^{het}	LC^{hom}	PRS^{het}	PRS^{hom}	AUT
β_1			0.6411^{***}		
	()		(0.0247)		
β_2			0.0006		
	()		(0.0016)		
β_3			0.0260		
	()		(0.1798)		
δ		0.9406^{***}			
	()	(0.0126)			
γ^2		0.1846^{***}	0.2104^{***}	0.1560^{***}	0.2287^{***}
	()	(0.0139)	(0.0073)	(0.0068)	(0.0094)
Log likelihood	−497.6	−999.3	−1078.3	−1227.9	−1508.8
R^2					
Observations	1468	1468	1468	1468	1468
	Vuong's tests				
LC^{hom}					
	()				
PRS^{het}		2.559^{***}			
	()	(0.005)			
PRS^{hom}		459.05^{***}	243.36^{***}		
	()	(0.002)	(0.000)		
AUT		1020.9^{***}	11.44^{***}	17.46^{***}	
	()	(0.000)	(0.000)	(0.000)	

Notes: LC^{het}: equation (25) is estimated by maximizing the partial log-likelihood function (27) by maximizing the log-likelihood function (27). LC^{hom}: equation (25), with $\beta = 0$, is estimated. PRS^{het}: equation (15) is estimated by maximizing the log-likelihood function (16). PRS^{hom}: equation (15), with $\beta = 0$ is estimated. AUT: equation (17) is estimated by maximizing the log-likelihood function (18). In the first panel standard errors are in parentheses. They have been calculated taken into account misspecification and serial correlation in the error terms. In the second panel, p-values of Vuong's tests are in parentheses, indicating whether the model of the line can be rejected to be as close to the true generating process as the model of the column. In the case of nested models, the p-values are simulated. * indicates significance at the 10% level, ** at 5%, and *** at 1%.

Table 3: Risk sharing *within villages* in Faisalabad district, adult-equivalent *nondurable* consumption

Model	LC^{het}	LC^{hom}	PRS^{het}	PRS^{hom}	AUT
β_1			0.3907***		
	()		(0.0350)		
β_2			−0.0000		
	()		(0.0026)		
β_3			0.3552		
	()		(0.3560)		
δ					
	()	()			
γ^2			0.1583***	0.1401***	0.2101***
	()	()	(0.0062)	(0.0061)	(0.0085)
Log likelihood			−1066.9	−1149.0	−1446.6
R^2					
Observations	1468	1468	1468	1468	1468
	Vuong's tests				
LC^{hom}					
	()				
PRS^{het}					
	()	()			
PRS^{hom}			158.11***		
	()	()	(0.000)		
AUT			14.05***	18.18***	
	()	()	(0.000)	(0.000)	

Notes: LC^{het}: equation (25) is estimated by maximizing the partial log-likelihood function (27) by maximizing the log-likelihood function (27). LC^{hom}: equation (25), with $\beta = 0$, is estimated. PRS^{het}: equation (15) is estimated by maximizing the log-likelihood function (16). PRS^{hom}: equation (15), with $\beta = 0$ is estimated. AUT: equation (17) is estimated by maximizing the log-likelihood function (18). In the first panel standard errors are in parentheses. They have been calculated taken into account misspecification and serial correlation. In the second panel, p-values of Vuong's tests are in parentheses, indicating whether the model of the line can be rejected to be as close to the true generating process as the model of the column. In the case of nested models, the p-values are simulated. * indicates significance at the 10% level, ** at 5%, and *** at 1%.

Tables 2 and 3 show that the coefficient of relative risk aversion is positively related to `education`, our proxy for wealth. This means that better educated, richer households bear less of the uninsured risk. This result is consistent with the hypothesis of Arrow (1965) and Pratt (1964) that relative risk aversion is increasing with wealth. `Age` and `gender` do not affect σ_i significantly, but the coefficient of `gender` has the expected positive sign.

(...)

5.2 Robustness checks

This section first looks at whether our results are robust to changing the number of simulations J in the limited commitment case. Then, we check for the effects of approximation error when computing the solution of the risk sharing with limited commitment model. In particular, we increase the number of gridpoints to 60 when discretizing the co-state variable x_i, for some households. Finally, we examine whether mismeasurement of income affects our main conclusions. To do this, we perturbate income multiplying each observation by $\exp(\varepsilon_{it}^y)$, where ε_{it}^y is a draw from $N(0, 0.4)$. Given a set of draws, we recompute all the estimates with $\tilde{y}_{it} = \exp(\varepsilon_{it}^y) y_{it}$ as our income data.

(...)

6 Policy simulations

In this section, we simulate the effects of a simple redistribution from the richest 30% of households in terms of mean income or wealth, to the bottom 30% percent. We look at three cases: (i) wealth transfers are made, (ii) income transfers are made, and (iii) consumption is transfered. The effects of the policies on poverty and inequality are discussed.

(...)

7 Extensions

In this section, we develop a more general empirical model. In particular, (i) we allow preferences to depend on time-varying household characteristics, (ii) preference shock are introduced, and (iii) income is measured with error as well as consumption. We argue that these extensions are feasible in theory, however, on the practical side, increased computation time is a serious concern in the case of risk sharing with limited commitment.

Instead of (1), let us now specify the utility function as

$$u_{it}(c_{it}) = \exp(\xi_{it}) \frac{c_{it}^{1-\sigma_{it}} - 1}{1 - \sigma_{it}}, \tag{29}$$

where

$$\xi_{it} = \eta_i + w_{it}'\alpha + \varepsilon_{it}^{\eta},$$

and the coefficient of relative risk aversion

$$\sigma_{it} = 1 + z_{it}'\beta,$$

where η_i is a time-invariant unobservable individual effect, as before, w_{it} and z_{it} are vectors of observable characteristics of household i at time t, α and β are parameter vectors to be estimated, and ε_{it}^{η} is a normally distributed preference shock with mean 0 and variance γ_{η}^2. Note that σ_{it} is only allowed to depend on observable covariates. Finally, let y_{it}^* be income observed by the econometrician. Measurement error in income is assumed to be multiplicative and log-normally distributed, that is,

$$y_{it}^* = \exp\left(\varepsilon_{it}^{y}\right) y_{it},$$

and $\varepsilon_{it}^{y} \sim N\left(0, \gamma_y^2\right)$.

7.1 Perfect risk sharing

The first order condition of household i, sharing risk with the rest of the community, can now be written as

$$\sigma_{it} \log c_{it} - \xi_{it} - \log c_{kt} = \log \lambda_i.$$

Replacing for ξ_{it} and σ_{it}, in terms of measured consumption we have

$$(1 + z_{it}'\beta) \log c_{it}^* - (1 + z_{it}'\beta)\, \varepsilon_{it} - \eta_i - w_{it}'\alpha - \varepsilon_{it}^{\eta} - \log c_{kt} = \log \lambda_i.$$

First differencing and rearranging gives

$$\log\left(\frac{c_{it}^*}{c_{kt}}\right) = \log\left(\frac{c_{i,t-1}^*}{c_{k,t-1}}\right) - z_{it}'\beta \log c_{it}^* + z_{i,t-1}'\beta \log c_{i,t-1}^* + w_{it}'\alpha - w_{i,t-1}'\alpha$$
$$+ \varepsilon_{it}^{\eta} - \varepsilon_{i,t-1}^{\eta} + (1 + z_{it}'\beta)\, \varepsilon_{it} - \left(1 + z_{i,t-1}'\beta\right) \varepsilon_{i,t-1}.$$

Thus now the error term is distributed as $N\left(0, 2\gamma_{\eta}^2 + (1 + z_{it}'\beta)^2 \gamma^2 + \left(1 + z_{i,t-1}'\beta\right)^2 \gamma^2\right)$.

7.2 Autarky

In the autarky case, preferences do not play a role, thus only the additional measurement error in income has to be taken into account. Instead of (17), the equation to be estimated is

$$\log\left(\frac{c_{it}^{*}}{c_{kt}}\right)=\log\left(\frac{c_{i,t-1}^{*}}{c_{k,t-1}}\right)+\log\left(\frac{y_{it}^{*}}{y_{kt}}\right)-\log\left(\frac{y_{i,t-1}^{*}}{y_{k,t-1}}\right)+\varepsilon_{it}-\varepsilon_{i,t-1}-\left(\varepsilon_{it}^{y}-\varepsilon_{i,t-1}^{y}\right),$$

and γ^2 and γ_y^2 are not jointly identified.

7.3 Risk sharing with limited commitment

Let us disregard measurement error for the moment. With the utility function (29), a typical enforcement constraint can be written as

$$\exp\left(w_{it}'\alpha+\varepsilon_{it}^{\eta}\right)\frac{\hat{c}_{it}\left(s^{t}\right)^{-z_{it}'\beta}-1}{-z_{it}'\beta}+\sum_{r=t+1}^{\infty}\sum_{s^{r}}\delta^{r-t}\pi\left(s^{r}\right)\exp\left(w_{ir}'\alpha+\varepsilon_{ir}^{\eta}\right)\frac{\hat{c}_{ir}\left(s^{r}\right)^{-z_{ir}'\beta}-1}{-z_{ir}'\beta}\geq$$
$$\geq\exp\left(w_{it}'\alpha+\varepsilon_{it}^{\eta}\right)\frac{y_{it}\left(s_{t}\right)^{-z_{it}'\beta}-1}{-z_{it}'\beta}+\sum_{r=t+1}^{\infty}\sum_{s_{r}}\delta^{r-t}\pi\left(s_{r}\right)\exp\left(w_{ir}'\alpha+\varepsilon_{ir}^{\eta}\right)\frac{y_{ir}\left(s_{r}\right)^{-z_{ir}'\beta}-1}{-z_{ir}'\beta}.$$

Note that future household characteristics and preference shocks enter into today's enforcement constraint. Therefore, some assumptions have to be made on households' expectations about their future characteristics and preferences shocks.

If we assume that households are myopic in the sense that they do not expect their characteristics to change relative to the rest of the village, further, they do not expect preference shocks to differ from today's, then we may divide the above equation by $\exp\left(w_{it}'\alpha+\varepsilon_{it}^{\eta}\right)=\exp\left(w_{ir}'\alpha+\varepsilon_{ir}^{\eta}\right),\forall r>t$. In this case computation time is only multiplied by the number of time periods. However, the assumption that households are always surprised by a change in any of their characteristics is very strong.

Alternatively, we may assume that households form rational expectations, and the expectations about their characteristics next period are the observed values. Preference shock can be integrated out using simulation, if we make some assumption about their distribution. To keep things tractable, we assume that preference shocks are i.i.d. over time and across households. Given household characteristics and a realization of the preference shock today, the enforcement constraint can be written in a recursive form as

$$\exp\left(w_{it}'\alpha + \varepsilon_{it}^{\eta}\right)\left[\frac{\hat{c}_i\left(s_t, x_{t-1}\right)^{-z_{it}'\beta} - 1}{-z_{it}'\beta} - \frac{y_i\left(s_t\right)^{-z_{it}'\beta} - 1}{-z_{it}'\beta}\right] \geq \delta \sum_{s_{t+1}} \pi\left(s_{t+1}\right) \times$$

$$\int \left[V_i^{aut}\left(s_{t+1}, w_{i,t+1}, z_{i,t+1}, \varepsilon_{i,t+1}^{\eta}\right) - V_i\left(s_{t+1}, x_t\left(s_t, x_{t-1}\right), w_{i,t+1}, z_{i,t+1}, \varepsilon_{i,t+1}^{\eta}\right)\right] f\left(\varepsilon_{i,t+1}^{\eta}\right) d\varepsilon_{i,t+1}^{\eta}.$$

Thus the model should be solved on an extended state space that includes household characteristics. In practice, a grid has to be defined over each characteristic. Today's preference shock also has to be integrated out by simulation. Computation time thus becomes prohibitively long. Note also that, with preference shocks but without measurement error in consumption, the equation to be estimated in the limited commitment case (25) becomes deterministic.

Finally, let us look at what changes if income is measured with error, as well as consumption. First, given the distribution from which income is drawn, measurement error does not affect the optimal intervals that characterize the constrained efficient risk sharing contract, since the model is solved using a grid on income. Second, the introduction of measurement error in income may plague the nonparametric estimation of the income process. We could perturbate observed income, and then recompute the model given a matrix of draws of the measurement error. Computation time is then proportional the number of matrices of draws, which would make the estimation take prohibitively long. Third, today's income observation directly affects consumption predicted by the model. Once again, simulation is a simple way to deal with this problem, and the solution of the model does not have to be recomputed, thus computation time increases only moderately.

8 Concluding remarks

This paper has performed statistical tests to distinguish between five models of risk sharing, namely, perfect risk sharing with and without preference heterogeneity, no risk sharing, or, autarky, and risk sharing with limited commitment with and without preference heterogeneity. Structural estimation results suggest that both heterogeneity in preferences and limitations to the enforcement of informal insurance contracts are important in explaining the consumption allocation in rural Pakistan.

Research on the structural modeling of how consumption is allocated between households in poor communities can serve as an input for policy design. It is important to take into account existing informal arrangements to share risk when evaluating redistributive policies or micro-insurance programs. Thus policy makers and members of non-governmental organizations could have a better understanding of the effects of their programs.

Several interesting extensions are possible. First, other models of risk sharing could be incorporated into the analysis, like the model of risk sharing with private information (Wang, 1995). Another important task for future work is to allow for savings, as in Ligon, Thomas, and Worrall (2000), instead of separating insurance across states from intertemporal consumption smoothing decisions, as in this paper. Third, when complete markets do not exist to insure against income fluctuations, households are expected to choose safer jobs, or safer production technologies in agriculture. In other words, they smooth income, not just consumption (Morduch, 1995). These ideas could be formalized in the context of this paper, endogenizing income by allowing households to choose between several income generating processes. Then, the cost of lack of insurance in terms of lower expected incomes could be quantified.

References

Ackerberg, D., J. Geweke, and J. Hahn (2008). Comments on "Convergence Properties of the Likelihood of Computed Dynamic Models" by Fernández-Villaverde, Rubio-Ramírez and Santos. *Econometrica*, forthcoming.

Akaike, H. (1973). Information Theory and an Extension of the Maximum Likelihood Principle. In B. N. Petrov and F. Csáki (Eds.), *Proceedings of the Second International Symposium of Information Theory*, pp. 257–281. Akadémiai Kiadó, Budapest.

Alderman, H. and M. Garcia (1993). Poverty, Household Food Security, and Nutrition in Rural Pakistan. Research report 96, International Food Policy Research Institute.

Arrow, K. (1965). *Aspects of the Theory of Risk-Bearing*. Yrjo Jahnssonin Saatio, Helsinki.

Attanasio, O. and J.-V. Ríos-Rull (2000). Consumption Smoothing in Island Economies: Can Public Insurance Reduce Welfare? *European Economic Review 44*(7), 1225–1258.

Binswanger, H. P. (1980). Attitude toward Risk: Experimental Measurement in Rural India. *American Journal of Agricultural Economics 62*(3), 395–407.

Blundell, R., L. Pistaferri, and I. Preston (2008). Consumption Inequality and Partial Insurance. *American Economic Review 98*(5), 1887–1921.

Borch, K. (1962). Equilibrium in a Reinsurance Market. *Econometrica 30*(3), 424–444.

Coate, S. and M. Ravallion (1993). Reciprocity without Commitment: Characterization and Performance of Informal Insurance Arrangements. *Journal of Development Economics 40*, 1–24.

Cochrane, J. H. (1991). A Simple Test of Consumption Insurance. *Journal of Political Economy 99*(5), 957–976.

Dercon, S. and P. Krishnan (2003a). Food Aid and Informal Insurance. UNU-WIDER Research Paper, World Institute for Development Economic Research.

Dercon, S. and P. Krishnan (2003b). Risk Sharing and Public Transfers. *Economic Journal 113*, 86–94.

Dixit, A., G. M. Grossman, and F. Gul (2000). The Dynamics of Political Compromise. *Journal of Political Economy 108*(3), 531–568.

Dubois, P. (2000). Assurance complète, hétérogénéité des préférences et métayage au Pakistan. *Annales d'Economie et de Statistique 59*, 1–36.

Dubois, P., B. Jullien, and T. Magnac (2008). Formal and Informal Risk Sharing in LDCs: Theory and Empirical Evidence. *Econometrica 76*(4), 679–725.

Fafchamps, M. (1999). Risk Sharing and Quasi-Credit. *Journal of International Trade and Economic Development 8*(3), 257–278.

Fernández-Villaverde, J., J. F. Rubio-Ramírez, and M. S. Santos (2006). Convergence Properties of the Likelihood of Computed Dynamic Models. *Econometrica 74*(1), 93–119.

Foster, A. D. and M. R. Rosenzweig (2001). Imperfect Commitment, Altruism, and the Family: Evidence from Transfer Behavior in Low-Income Rural Areas. *Review of Economics and Statistics 83*(3), 389–407.

Gouriéroux, C. and A. Monfort (1997). *Simulation-based Econometric Methods.* Oxford University Press, Oxford.

Gourinchas, P.-O. and J. A. Parker (2002). Consumption over the Life Cycle. *Econometrica 70*(1), 47–89.

Grimard, F. (1997). Household Consumption Smoothing Through Ethnic Ties: Evidence from Côte d'Ivoire. *Journal of Development Economics 53*(2), 391–422.

Guiso, L. and M. Paiella (2008). Risk Aversion, Wealth, and Background Risk. *Journal of the European Economic Association 6*(6), 1109–1150.

Halek, M. and J. G. Eisenhauer (2001). Demography of Risk Aversion. *Journal of Risk and Insurance 68*(1), 1–24.

Holt, C. A. and S. K. Laury (2002). Risk Aversion and Incentive Effects. *American Economic Review 92*(5), 1644–1655.

Jianakoplos, N. A. and A. Bernasek (1998). Are Women More Risk Averse? *Economic Inquiry 36*(4), 620–630.

Kehoe, P. J. and F. Perri (2002). International Business Cycles with Endogenous Incomplete Markets. *Econometrica 70*(3), 907–928.

Kocherlakota, N. R. (1996). Implications of Efficient Risk Sharing without Commitment. *Review of Economic Studies 63*(4), 595–609.

Krueger, D. and F. Perri (2006). Does Income Inequality Lead to Consumption Inequality? Evidence and Theory. *Review of Economic Studies 73*(1), 163–193.

Kurosaki, T. and M. Fafchamps (2002). Insurance Market Efficiency and Crop Choices in Pakistan. *Journal of Development Economics 67*(2), 419–453.

Ligon, E., J. P. Thomas, and T. Worrall (2002). Informal Insurance Arrangements with Limited Commitment: Theory and Evidence from Village Economies. *Review of Economic Studies 69*(1), 209–244.

Marcet, A. and R. Marimon (1998). Recursive Contracts. EUI ECO 1998/37, European University Institute.

Mazzocco, M. (2007). Household Intertemporal Behaviour: A Collective Characterization and a Test of Commitment. *Review of Economic Studies 74*(3), 857–895.

Mazzocco, M. and S. Saini (2008). Testing Efficient Risk Sharing with Heterogeneous Risk Preferences.

Morduch, J. (1995). Income Smoothing and Consumption Smoothing. *Journal of Economic Perspectives 9*(3), 103–114.

Ogaki, M. and Q. Zhang (2001). Decreasing Relative Risk Aversion and Tests of Risk Sharing. *Econometrica 69*(2), 515–526.

Palsson, A.-M. (1996). Does the Degree of Relative Risk Aversion Vary with Household Characteristics? *Journal of Economic Psychology 17*(6), 771–787.

Pratt, J. W. (1964). Risk Aversion in the Small and in the Large. *Econometrica 32*(1/2), 122–136.

Renkow, M. (1993). Land Prices, Land Rents, and Technological Change: Evidence from Pakistan. *World Development 21*(5), 791–803.

Rivers, D. and Q. H. Vuong (2002). Model Selection Tests for Nonlinear Dynamic Models. *Econometrics Journal 5*(1), 1–39.

Schechter, L. (2007). Theft, Gift-Giving, and Trustworthiness: Honesty Is Its Own Reward in Rural Paraguay. *American Economic Review 97*(5), 1560–1582.

Schulhofer-Wohl, S. (2007). Heterogeneity, Risk Sharing and the Welfare Costs of Idiosyncratic Risk.

Shaw, K. L. (1996). An Empirical Analysis of Risk Aversion and Income Growth. *Journal of Labor Economics 14*(4), 626–653.

Thomas, J. and T. Worrall (1988). Self-Enforcing Wage Contracts. *Review of Economic Studies 55*(4), 541–554.

Townsend, R. M. (1994). Risk and Insurance in Village India. *Econometrica 62*(3), 539–591.

Vuong, Q. H. (1989). Likelihood Ratio Tests for Model Selection and Non-nested Hypotheses. *Econometrica 57*(2), 307–333.

Wang, C. (1995). Dynamic Insurance with Private Information and Balanced Budgets. *Review of Economic Studies 62*(4), 577–595.

Wang, H. and S. D. Hanna (1997). Does Risk Tolerance Decrease with Age? *Financial Counseling and Planning 8*(2), 27–32.

Wilson, R. (1968). The Theory of Syndicates. *Econometrica 36*(1), 119–132.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data.* MIT Press, Cambridge, Massachusetts.

Appendix

This appendix details how to solve the model of risk sharing with limited commitment numerically. The aim is to compute the consumption predicted by the model, $\hat{c}_{it}$, relative to mean consumption in the community, for each household i at each time t, taking preference parameters and the discrete distributions from which incomes are drawn as given.[13]

The model is solved for each household separately, since we take observables entering the utility function, as well as the household specific Y_i, as given. Consider some household i sharing risk with the rest of the community, represented by household k. Take preference parameters β, the discount factor δ, the variance of measurement error γ^2, and the discrete distributions from which incomes are drawn F_{Y_i} and F_{Y_k} as given. The aim is to solve for the decision variables, that is, (true) consumption $c_{it}(s_t, x_{i,t-1})$, the (true) relative weight of household i $x_{it}(s_t, x_{i,t-1})$, and for the lifetime utility household i gets from her consumption stream being in the informal risk sharing arrangement $V_i(s_t, x_{i,t-1})$, given the state of the world today (s_t, x_{t-1}).

The first task is to solve for the optimal intervals characterizing the solution of the model. Define a grid over the continuous variable x_i for each value of s_t (we define the same points for all s_t). The support of the grid is the range of ratios of marginal utilities of household k and i given the income and consumption observations. We define an equidistant grid on $\log x_i$ of 30 points. Guess a solution for the value functions, that is, guess $V_i^0(s_t, x_{i,t-1})$, for each gridpoint. Unfortunately, the algorithm does not converge from any initial guess for the value functions, but the value of the perfect risk sharing case will do.[14]

Now proceed to update the guess. Suppose we are at the h^{th} iteration. Let us look at gridpoint $(\tilde{s}_t, \tilde{x}_{i,t-1})$. Three cases have to be distinguished: (a) neither enforcement con-

[13] The first step of estimation involves determining these distributions (see main text), while the last step is the maximization over the remaining parameters, which is done using a standard optimization algorithm available in R (function "optim()" with method "L-BFGS-B"). Now we talk about the computation in between these steps.

[14] Characterizing the convergence properties of the algorithm is left for future research. However, we know that the algorithm does not converge to the constrained-efficient solution from any initial guess for the value functions. For example, if we set the initial guess $V_i^0(s_t, x_{i,t-1})$ equal to the autarkic values, every iteration yields these same autarkic values. This is natural, since autarky is also a subgame perfect Nash equilibrium (SPNE).

straint binds, (b) the enforcement constraint for household i binds, and (c) the enforcement constraint for household k binds. Note that the two enforcement constraints cannot bind at the same time, because only one of the two households may be called upon to make a positive net transfer.

We first suppose that the enforcement constraints do not bind, that is, we try to keep x_i constant. This means setting $\hat{x}_{it}^h = \tilde{x}_{i,t-1}$ at state $\tilde{s}_t$, where the upper index h refers to iteration h. Then, using the first order condition and the resource constraint, we get the consumption allocation $\left(\hat{c}_{it}^h, \hat{c}_{kt}^h\right)$. Now the enforcement constraints have to be checked. This means verifying whether

$$\frac{\left(\hat{c}_{it}^h\right)^{-z_i'\beta} - 1}{-z_i'\beta} + \delta \sum_{s_{t+1}} \pi\left(s_{t+1}\right) V_i^{h-1}\left(s_{t+1}, \hat{x}_{it}^h\right) \geq U_i^{aut}\left(\tilde{s}_t\right) \tag{30}$$

and

$$\log \hat{c}_{kt}^h + \delta \sum_{s_{t+1}} \pi\left(s_{t+1}\right) V_k^{h-1}\left(s_{t+1}, \hat{x}_{kt}^h\right) \geq U_k^{aut}\left(\tilde{s}_t\right). \tag{31}$$

Note the upper index $h-1$ for V_i and V_k, that is, we use the value function from the previous iteration.

(a) *The enforcement constraints (30) and (31) do not bind.* This is the easy case, since we have already computed $\hat{x}_{it}^h$ and the consumption allocation assuming that the enforcement constraints do not bind. What remains to be done is to set

$$V_i^h\left(\tilde{s}_t, \tilde{x}_{i,t-1}\right) = \frac{\left(\hat{c}_{it}^h\right)^{-z_i'\beta} - 1}{-z_i'\beta} + \delta \sum_{s_{t+1}} \pi\left(s_{t+1}\right) V_i^{h-1}\left(s_{t+1}, \hat{x}_{it}^h\right)$$

and

$$V_k^h\left(\tilde{s}_t, \tilde{x}_{i,t-1}\right) = \log \hat{c}_{kt}^h + \delta \sum_{s_{t+1}} \pi\left(s_{t+1}\right) V_k^{h-1}\left(s_{t+1}, \hat{x}_{it}^h\right)$$

(b) *The enforcement constraint (30) is binding.* In theory, we may compute $\hat{c}_{it}^h$ and $\hat{x}_{it}^h$ using (30) with equality and the first order condition. However, we do not know $V_i^{h-1}\left(s_{t+1}, \hat{x}_{it}^h\right)$ for any value of $\hat{x}_{it}^h$, only for the points on the grid. Therefore, we look for the point $\hat{x}_{it}^h$ for which $\hat{c}_{it}^h$ is such that (30) is satisfied and the left hand side is closer to the right hand side than for any other gridpoint. Finally, we update the value function as in case (a).

(c) *The enforcement constraint (31) is binding.* We proceed similarly as in case (b).

Now we are done with gridpoint $(\tilde{s}_t, \tilde{x}_{t-1})$. We have to do the above steps at all other gridpoints as well. Then the h^{th} iteration is complete. We keep iterating until the policy function converges, that is, the optimal intervals on x_{it} do not change. More precisely, the solution has been found, if the length of the difference between the endpoints of the optimal intervals at iteration $h-1$ and h is less than 0.001. In practice we allow for maximum 20 iterations. At the end, we have the the solution in the form $[\underline{x}_{it}(s_t, x_{i,t-1}), \overline{x}_{it}(s_t, x_{i,t-1})]$.

Computing the consumption of household i at time t, relative to mean consumption in the community, as predicted by the model is then done as follows. Remember that $c^*_{i,t-1}$ is observed consumption of household i at time $t-1$, and $\varepsilon^j_{i,t-1}$ is a realization of measurement error drawn from $N(0, \gamma^2)$. We compute $x_{i,t-1} = \left(\exp\left(-\varepsilon^j_{i,t-1}\right) c^*_{i,t-1}\right)^{1+z_i'\beta} / c_{k,t-1}$, and check whether it is in the optimal interval for today's state s_t. Since only a discrete number of income states have been considered, we map observed incomes into the income states of the model by picking the closest point for each household. We have to consider the above three cases.

(a) If $x_{i,t-1} \in [\underline{x}_{it}(s_t, x_{i,t-1}), \overline{x}_{it}(s_t, x_{i,t-1})]$, then we set $x_{it} = x_{i,t-1}$.

(b) If $x_{i,t-1} < \underline{x}_{it}(s_t, x_{i,t-1})$, then we determine it from (30) with equality, using a linear interpolation of the value functions from the last iteration over the gridpoints $\tilde{x}^1_{t-1}$ and $\tilde{x}^2_{t-1}$ closest to $x_{i,t-1}$ such that $\tilde{x}^1_{t-1} < x_{i,t-1} < \tilde{x}^2_{t-1}$. Thereby we make the predicted consumption allocation a continuous function of the parameters to be estimated at the end.

(c) If $x_{i,t-1} < \overline{x}_{it}(s_t, x_{i,t-1})$, then we use (31) with equality, and proceed similarly as in case (b).

Then we may write the likelihood of observation it, given $\varepsilon^j_{i,t-1}$, by plugging the $\hat{c}_{it}$ and $\hat{c}_{kt}$ computed here into (26).