RISK SHARING WITH LIMITED COMMITMENT AND PREFERENCE HETEROGENEITY: STRUCTURAL ESTIMATION AND TESTING

Sarolta Laczó
Institut d'Anàlisi Econòmica (IAE-CSIC)
and Barcelona GSE

Abstract
In order to analyze the role of limited commitment and preference heterogeneity in explaining the consumption allocation, I propose a theoretical and empirical framework to estimate and evaluate a risk-sharing model where insurance transfers have to be self-enforcing and the coefficient of relative risk aversion may depend on observable household characteristics. I compare this model to benchmark models with full commitment and/or without preference heterogeneity using data from three Indian villages. I find that the limited commitment model with heterogeneous preferences outperforms the benchmark models in a statistical sense and in terms of (i) explaining the dynamic response of consumption to idiosyncratic income shocks, (ii) accounting for the variation of consumption unexplained by household and time effects, and (iii) capturing the variation of inequality across time and villages and predicting changes in inequality. I also use the estimated models to predict the effects of a counterfactual tax and transfer policy on the consumption allocation. The limited commitment model with preference heterogeneity predicts larger benefits to the poor than its homogeneous counterpart. (JEL: C52, D10, D52)

1. Introduction

Households living in rural areas of low-income countries face a great amount of risk. Revenue from agricultural production is usually low and volatile as a result

DOI: 10.1111/jeea.12115

of extreme weather conditions, such as erratic monsoon rains in South Asia. Further, outside job opportunities are often lacking, and access to financial instruments to insure against consumption fluctuations is limited. In such an environment, households in a community rely on one another for insurance.

There is ample empirical evidence that households in poor villages achieve a remarkable amount of insurance, but they do not fully share the risks they face (Townsend 1994, and many others). Moreover, direct evidence shows that households make state-contingent transfers to one another (Udry 1994; Kinnan 2010). The literature has focused on two imperfections to explain the observed partial insurance: private information (Wang 1995; Ligon 1998) and lack of commitment (Kocherlakota 1996; Ligon, Thomas, and Worrall 2002; LTW hereafter). This paper focuses on the case where limited commitment (LC hereafter) is the friction that may cause a deviation from perfect risk sharing. Households in small communities can often observe shocks faced by their neighbors (such as bad harvest or illness), but no authority exists to enforce informal risk-sharing contracts. The approach of this paper could also be used to compare models of mutual insurance with private information as an alternative or additional friction. However, this is left for future research.

In addition to studying the relevance of LC, I investigate whether heterogeneity in risk preferences is useful in explaining the consumption allocation. In particular, I allow households' coefficients of relative risk aversion to depend on observable household characteristics. This is an important extension, because efficient risk sharing has two main implications. First, incomes should be pooled. Second, less risk-averse households should bear more uninsurable risk (Borch 1962; Wilson 1968). Assuming that risk preferences are homogeneous, one excludes an additional motive for risk sharing.

I show under what conditions the parameters of the LC model are identified. Additional parameters compared to the perfect risk-sharing model are pinned down by binding enforcement/participation constraints. I derive simulated maximum likelihood estimators taking into account measurement error in income and consumption and an unobservable heterogeneity term in the curvature of the utility function. I solve the LC model on a grid of parameters by standard value function iteration, but exploiting the characteristics of its solution at each step. At the estimation stage I interpolate the solution, hence reducing computational time. I statistically compare risk-sharing models in terms explaining consumption shares given income shares, using likelihood ratio-based tests introduced by Vuong (1989).

I apply the estimation and model selection methods to study risk sharing in three Indian villages, using data collected by the International Crop Research Institute for the Semi-Arid Tropics (ICRISAT). I find that the risk-sharing model with both LC and preference heterogeneity fits the consumption data from each village significantly better in a statistical sense than models with full commitment and/or without preference heterogeneity. Further, I find that the LC model with heterogeneous preferences outperforms the benchmark models in terms of (i) explaining the response of consumption to idiosyncratic income shocks, (ii) accounting for the variation of consumption unexplained by household and time effects, and (iii) capturing the

variation of inequality across time and villages and predicting changes in inequality, as measured by the variance of log consumption and the Gini coefficient.

Finally, I simulate the effects on the consumption allocation of a counterfactual tax and transfer policy, which can be thought of as public insurance (Krueger and Perri 2011; Broer 2011),[1] taking into account existing informal risk-sharing arrangements. According to perfect risk-sharing models, such a policy has no effect on the consumption allocation, while LC models are able to predict a redistribution of consumption. The LC model with preference heterogeneity predicts larger benefits to the poor than its homogeneous counterpart. The policy simulation exercise illustrates the quantitative importance of LC and heterogeneity in risk preferences for ex-ante policy evaluation (Todd and Wolpin 2006, 2008).

The rest of the paper is structured as follows. Next, I discuss the related literature. Section 2 details the theoretical models of risk sharing. In Section 3, the empirical models are set up and simulated maximum likelihood estimators are derived. Section 4 presents the application to household survey data from three Indian villages, including the counterfactual policy simulation. Section 5 concludes.

Related Literature

The early literature on risk sharing in village economies includes Townsend (1994), Grimard (1997), Dubois (2000), Dercon and Krishnan (2003a, b), and others. The tests of perfect risk sharing performed by these papers are of reduced form and alternative models of risk sharing are not examined.[2] These papers reject perfect risk sharing, but find that only a small fraction of idiosyncratic income shocks translates into consumption fluctuations.

A next wave of the literature explicitly considers alternative models of risk sharing, with frictions such as LC and asymmetric information, but only studies their reduced-form implications, see Fafchamps (1999), Attanasio and Ríos-Rull (2000), Foster and Rosenzweig (2001), Dubois, Jullien, and Magnac (2008), and Kinnan (2010). Kinnan (2010) looks at LC as well as hidden income and effort as potential impediments to risk sharing in Thai villages. She is able to reject some implications of all models except for the hidden income one.[3]

---

1.   See also Attanasio and Ríos-Rull (2000), who argue that, under LC, formal insurance provided by the state may crowd out informal insurance transfers to the extent that welfare decreases. They then provide reduced-form evidence on the crowding out of informal transfers as a result of the Progresa program in Mexico, but do not use the model to predict the transfers.

2.   See also the seminal papers by Cochrane (1991) and Mace (1991) for tests of perfect risk sharing in the United States.

3.   However, the inverse Euler equation, which she rejects, should hold for the hidden income model as well, not just for hidden effort; see Ligon (1998). Further, the hidden-income model can be written as a special case of the hidden-effort model: specify the production function as $y = e$, where y is output and e is effort, and assume that the cost of effort is linear.

The first paper which acknowledges heterogeneity in risk preferences when testing perfect risk sharing is Altug and Miller (1990). They allow preferences to depend on household demographics. Dubois (2000) provides a more powerful, nondirectional test. Mazzocco and Saini (2012) construct nonparametric tests of perfect risk sharing allowing for preference heterogeneity, and reject perfect risk sharing within villages but not within caste groups in Indian villages. Chiappori et al. (2013) estimate a coefficient of relative risk aversion for each household under the assumption that perfect risk sharing occurs, using an 84-wave panel from Thailand. They find substantial heterogeneity in risk preferences and that consumption insurance is close to perfect. I test perfect risk sharing against a well-specified alternative while allowing for heterogeneity in risk preferences.

LTW estimate the model of risk sharing with LC in a structural manner. However, they do not perform any statistical tests on parameters or model selection. In a recent paper, Karaivanov and Townsend (2014), building on Paulson, Townsend and Karaivanov (2006), apply Vuong's test, like the present paper, to compare risk-sharing models with different frictions—in particular, they consider models with asymmetric information as well. In the rural part of their Thai sample and with income and consumption data alone, Karaivanov and Townsend reject full risk sharing, but not by much, and find that the moral hazard regime fits the data best, sometimes statistically tied with LC or savings only. The main contribution of the present paper with respect to their work is to introduce heterogeneity in risk preferences. Their modeling and solution strategy also differ in that they assume the presence of a financial intermediary, while I consider a model of mutual insurance among households as LTW.

Karaivanov (2012) applies a similar methodology to study which type of financial friction (savings only, borrowing and lending with default, or moral hazard constrained insurance) is most useful in explaining the choice to become an entrepreneur in Thailand. Relatedly, Schulhofer-Wohl (2011) uses an experimental measure of risk aversion and finds evidence that occupational choice is affected by risk preferences in the United States. 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. Accounting for how risk preferences and/or lack of consumption insurance affect occupational choice and/or the choice of production technologies is beyond the scope of the present paper.

This paper is also related to the literature on explaining changes over time in consumption inequality given income inequality. Krueger and Perri (2006) find that qualitatively LC can account for consumption inequality increasing less than income inequality in the United States over the period 1980–2003. However, it implies too much risk sharing—that is, not enough increase in consumption inequality. On the other hand, Blundell, Pistaferri and Preston (2008) document that income shocks have become less persistent, hence easier to insure against. Structural estimation results similar to the ones presented in this paper could be used to predict the effects of such changes in income processes on consumption inequality.

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 its expected lifetime utility,

$$\mathbb{E}_0 \sum_{t=1}^{\infty} \delta^t u_i(c_{it}),$$

where $\mathbb{E}_0$ is the expected value at time 0 calculated with respect to the probability measure describing the common beliefs about income processes, $\delta \in (0, 1)$ is the (common) discount factor, $u_i()$ describes the instantaneous preferences of household i and is strictly increasing and strictly concave, and c_{it} is consumption by household i at time t.

Each household i is endowed with random income Y_i. Y_i follows a Markov process and is independent across households. The distribution of Y_i, $\forall i$, is common knowledge ex ante, and so are income realizations ex post at each time t. That is, there are no informational problems. Let s_t denote the state of the world which describes the income realizations of all households in the community at time t, and s^t denote the history of income states, that is, $s^t = (s_1, \ldots, s_t)$. Note that income is exogenous by assumption. That is, the effect of risk on choices among different income-generating processes is ignored.

I interpret the models as predicting consumption shares, given income shares and aggregate income/consumption in the community.[4] In other words, any difference between household consumption and income is thought of as a transfer to or from the rest of the village, and not as saving or dissaving explicitly.

The rest of this section describes two models in turn. First, I consider the model of perfect risk sharing. Second, I detail the model of risk sharing with LC.

2.1. Perfect Risk Sharing

To find the Pareto-optimal allocations, one can solve the following problem: the (utilitarian) social planner maximizes a weighted sum of households' expected lifetime utilities,

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

where λ_i is the (initial) Pareto weight of household i in the social planner's objective, $\pi(s^t)$ is the probability of history s^t occurring, and $c_{it}(s^t)$ denotes the consumption of household i when history s^t has occurred, subject to the resource constraint

$$\sum_i c_{it}(s^t) \leq \sum_i y_{it}(s_t), \quad \forall s^t, \forall t, \tag{2}$$

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

---

4.   Income from the data is rescaled so that aggregate consumption and aggregate income are always equal.

The well-known result that for any two households i and v,

$$\frac{u'_v(c_{vt}(s^t))}{u'_i(c_{it}(s^t))} = \frac{\lambda_i}{\lambda_v}, \quad \forall s^t, \forall t, \tag{3}$$

that is, the ratio of marginal utilities 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). Equation (3) implies that all idiosyncratic risks are insured away, and households share aggregate risk efficiently. In particular, less risk-averse households bear more uninsurable risk.

2.2. Risk Sharing with Limited Commitment

To find the constrained-efficient consumption allocations, I follow Kehoe and Perri (2002), but consider an endowment economy, and solve the following problem: the social planner maximizes equation (1), subject to equation (2) and the enforcement/participation constraints (PCs),

$$\sum_{r=t}^{\infty} \sum_{s^r} \delta^{r-t} \pi(s^r \mid s^t) u_i(c_{ir}(s^r)) \geq U_i^{aut}(s_t), \quad \forall s^t, \forall t, \forall i, \tag{4}$$

where $(s^r \mid s^t)$ is the probability of history s^r occurring given that history s^t occurred up to time t. The right-hand side of (4) is the value of autarky, namely the outside option, for household i at state s_t and time t. LC means that each household may deviate and revert to autarky upon receiving its current income. Hence, the mutual insurance contract must provide at least as much lifetime utility as autarky given any history of income shocks. Note that the PCs, (4), assume that if a household deviates, other households in the community do not enter into any risk-sharing arrangement with it in the future. That is, households apply a grim trigger strategy.

In autarky households consume their own income, and the community may impose additional punishments on defectors, such as exclusion from social activities, as in LTW. For ease of interpretation, I model such punishments as a fraction of consumption each period. Hence, in autarky $c_{it}(s_t) = (1 - \varphi) y_{it}(s_t), \forall s_t, \forall t, \forall i,$ where $0 \leq \varphi < 1$. Then, the value of autarky can be computed by iterating the Bellman equation

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

Individual savings are assumed absent, as in Kocherlakota (1996), LTW, and others.[5]

---

5.   One could potentially assume that households in autarky have access to a storage technology. While a solution can often be found for a given set of parameter values, the non-emptiness of the feasible set is not guaranteed, because for some discount factors some households would be better off in autarky than in a risk-sharing arrangement without storage in equilibrium. Then, aggregate consumption becomes

The problem is not recursive, because future decision variables enter into today's PCs. Therefore, even if income were i.i.d., consumption may depend on the whole history of income realizations.

Denoting the multiplier on the PC of household i, equation (4), by $\delta^t \pi(s^t) \mu_i(s^t)$, and the multiplier on the resource constraint, equation (2), by $\delta^t \pi(s^t) \rho(s^t)$, the Lagrangian is

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

Using the ideas of Marcet and Marimon (2011), the Lagrangian can be written in the form

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

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 current Pareto weight of household i, and it is equal to its initial Pareto weight plus the sum of the Lagrange multipliers on its PCs along the history s^t.

Consider household i sharing risk with household v again. Define

$$x_i(s^t) \equiv \frac{M_i(s^t)}{M_v(s^t)},$$

the relative Pareto weight assigned to household i when history s^t has occurred, normalizing the weight of household v to 1. Think of household v as the average household in a village, hence the normalization is $(1/N) \sum_i x_i(s^t) = 1$, $\forall s^t$.

The vector of relative weights $x(s^t)$, with elements $x_i(s^t)$, can be used as a co-state variable. The solution consists of the policy functions $x_{it}(s_t, x_{t-1})$ and $c_{it}(s_t, x_{t-1})$,

---

endogenous. Further, it is unclear why such a technology would not be available in equilibrium as well, if it were available in autarky. Ábrahám and Laczó (2013) analyze the LC model with storage, including in equilibrium and at the aggregate level. Estimating that model is computationally challenging, further, (public) assets are not observed in the ICRISAT data set used in this paper.

$\forall i$, where x_{t-1} is a sufficient statistic for everything that happened in the past. The optimality condition which links consumptions and the current relative Pareto weights is

$$\frac{u'_v(c_{vt}(s_t, x_{t-1}))}{u'_i(c_{it}(s_t, x_{t-1}))} = x_{it}(s_t, x_{t-1}), \quad \forall i. \tag{6}$$

At last, the value functions can be written recursively as

$$V_i(s_t, x_{t-1}) = u_i(c_{it}(s_t, x_{t-1})) + \delta \sum_{s_{t+1}} \pi(s_{t+1}) V_i(s_{t+1}, x_t(s_t, x_{t-1})). \tag{7}$$

The solution is fully characterized by a set of state-dependent optimal intervals (LTW), which give the relative Pareto weights for which both PCs are satisfied in each income state s for household i. Denote the interval for household i sharing risk with the "rest of the village" for state s by $[\underline{x}_i^s, \bar{x}_i^s]$. Online Appendix A describes how the intervals are computed.

I then use only $\underline{x}_i^s$, which is pinned down by the PC of household i, $\forall s$, $\forall i$, to find the predicted consumption allocation. However, the values for $\underline{x}_i^s$ are only approximations of the values of the N-household case, because tomorrow's value functions depend on $\bar{x}_i^s$, $\forall s$, which are determined by PCs binding for the average household, v. That is, I assume that the distribution of incomes (and hence of the Pareto weights) and preference heterogeneity for the rest of the village is not of first order. Instead, each of the other $N-1$ households receives the village mean income and has average risk aversion. Note that the idea is similar to that of Krusell and Smith (1998).[6] Note also that the quality of approximation may differ across models, and the model comparisons will be made given the unknown quality of approximation.

Given $\underline{x}_i^s$, $\forall s$, $\forall i$, an iterated version of the updating rule of LTW can be used to compute the predicted relative Pareto weights at time t, $\hat{x}_t$, and the predicted consumptions, $\hat{c}_t$; given x_{t-1}, computed from the consumption allocation of time $t-1$ (see what follows), and $s_{it} = (y_{it}, y_{vt})$, $\forall i$, where y_{it} is the closest grid point to household i's income at time t in the data and y_{vt} is the closest grid point to average income in the village at time t.

1. For each household, check whether $x_{i,t-1} < \underline{x}_i^{s_{it}}$. If so, set $\hat{x}_{it} = \underline{x}_i^{s_{it}}$ and compute $\hat{c}_{it}$ from equation (6) with c_{vt} equal to average consumption of the rest of the village in state s_{it}—that is, to $(y_{it} + (N-1)y_{vt} - \hat{c}_{it})/(N-1)$.

2. For the remaining households—that is, $\forall i$ such that $x_{i,t-1} \geq \underline{x}_i^{s_{it}}$—set $\hat{x}_{it} = x_{i,t-1}$ and compute $\hat{c}_{it}$ from (6) now with c_{vt} equal to average consumption of all

---

6. Solving the model with N households would require $N + (N-1)$ state variables (each household's income and the relative Pareto weights) with $N \geq 31$ in the application, which is infeasible. Both LTW and Dubois, Jullien, and Magnac (2008) use a similar household–rest-of-the-village characterization. LTW have used this approach to approximate the solution to a simple example with many (500) households, which should be quite close to the continuum case. They find that the average value of the correlation coefficient between the two consumption paths is 0.972.

unconstrained households. Then compute the "real" $\hat{x}_{it}$ using the left-hand side of equation (6) given $\hat{c}_{it}$ just computed and c_{vt} equal to average consumption of the rest of the village. Set $x_{i,t-1} = x_{it}$.

3. Repeat steps 1 and 2 for all unconstrained households of the previous iteration and with the fictional $x_{i,t-1}$ computed in step 2. Repeat until no PC is violated.

4. Set $\hat{c}_{vt,i} = (y_{it} + (N-1)y_{vt})/N$, which represents average income and consumption in the village from the model.

3. Empirical Models

Assume that the utility function of each household i takes the isoelastic (CRRA) form, namely

$$u_i(c_{it}) = \frac{c_{it}^{1-\sigma_i} - 1}{1 - \sigma_i}, \tag{8}$$

where $\sigma_i > 0$ is the coefficient of relative risk aversion (RRA) of household i. Let σ denote average risk aversion in the village (and the risk aversion of the average household v), namely $\sum_{i=1}^{N} \sigma_i / N = \sigma$. I allow σ_i to depend on time-invariant observable covariates of household i, denoted z_i, with elements $z_{ik}, k \in \{1, 2, \ldots, K\}$, normalized so that $\mathbb{E}z_{ik} = 0$, $\forall k$, in each village, and on an unobservable term. That is,

$$\sigma_i \sim \log N\left(\mu_i, \gamma_\sigma^2\right) \qquad \text{with} \qquad \mu_i = \log\left(\sigma + z_i'\beta\right) - \frac{\gamma_\sigma^2}{2},$$

where β is a parameter vector to be estimated, with elements β_k. Note that $\mathbb{E}\sigma_i = \sigma + z_i'\beta$. Let $\sigma_i^* = \sigma + z_i'\beta$ and $\varepsilon_i^\sigma = \sigma_i - \sigma_i^*$. The term $z_i'\beta$ captures heterogeneity in the curvature of the utility function across households which can be related to observables, while ε_i^σ captures unexplained heterogeneity.[7]

I allow for measurement error in consumption and income, and assume that they are multiplicative and log-normally distributed. Let c_{it}^* denote consumption observed by the econometrician, and let $\exp(\varepsilon_{it}^c)$ be the multiplicative measurement error in household i's consumption at time t. Then, we may write $c_{it}^* = \exp(\varepsilon_{it}^c)c_{it}$, where ε_{it}^c is i.i.d. across households and time, and $\varepsilon_{it}^c \sim N(0, \gamma_c^2)$, where γ_c^2 is to be estimated.[8] The measurement error in income, denoted ε_{it}^y, has the same properties, and its variance is denoted by γ_y^2. Note that true consumption, c_{it}, and true income, y_{it}, are assumed to be observed by all households in the community.

I model the allocation of observed consumption $c_t^* \equiv (c_{1t}^*, \ldots, c_{it}^*, \ldots, c_{Nt}^*)$, for $t = 2, \ldots, T$, given observed consumption at time 1, c_1^*, the history of observed

---

7.   Guiso and Paiella (2008) construct a direct measure of risk aversion using the 1995 Bank of Italy Survey of Household Income and Wealth and find that a substantial part of the heterogeneity in risk preferences across households cannot be explained by observable household characteristics.

8.   Measurement error in consumption accounts for the error term in the estimating equations that follow.

income realizations, y_t^*, for $t = 2, \ldots, T$, time-invariant household characteristics, $Z = [z_1, \ldots, z_i, \ldots, z_N]'$, and parameters. In mathematical terms, I derive how the following conditional density can be specified based on the above models of risk sharing:

$$f\left(c_T^*, \ldots, c_2^* \mid c_1^*, y_T^*, \ldots, y_2^*, \lambda, Z; \sigma, \beta, \delta, \varphi, \gamma_c^2, \gamma_y^2, \gamma_\sigma^2, F_Y\right), \qquad (9)$$

where the vector of Pareto weights, λ, is a nuisance parameter, $\theta \equiv (\sigma, \beta, \delta, \varphi, \gamma_c^2, \gamma_y^2)$ are the parameters to be estimated,[9] I consider $\gamma_\sigma^2 = 0$ as the benchmark (i.e., no unobserved heterogeneity in the curvature of the utility function) and $\gamma_\sigma^2 = 0.02$ as a robustness check, and F_Y summarizes households' income processes. Each of the models of risk sharing of Section 2 allows us to factorize the density (9). In particular, we may write

$$\prod_{t=2,\ldots,T} f\left(c_t^* \mid y_t^*, x_{t-1}, Z; \theta, \gamma_\sigma^2, F_Y\right), \qquad (10)$$

where x_{t-1} is the vector of relative Pareto weights at time $t-1$, which has elements $x_{i,t-1}$ and is not observed (I deal with this issue in what follows). Note that past incomes and consumptions only matter through x_{t-1}. Further, remember that current aggregate income and consumption are assumed given. The models explain consumption shares, but not the evolution of aggregate consumption.

Given the current relative Pareto weights, x_t, one can write the optimality conditions as

$$\frac{u_v'(c_{vt})}{u_i'(c_{it})} = \frac{(c_{vt})^{-\sigma}}{(c_{it})^{-\sigma_i}} = x_{it}, \quad \forall i. \qquad (11)$$

The next two sections detail in turn what the model of perfect risk sharing (Section 3.1) and risk sharing with LC (Section 3.2) say about x_{it} in equation (11), show which parameters are identified, and discuss how the models are estimated using simulated (pseudo) maximum likelihood. Afterwards, Section 3.3 expands on the model selection test I apply, namely, Vuong's general test.

3.1. Perfect Risk Sharing

In the case of perfect risk sharing, the current consumption allocation depends only on current exogenous variables and not on past ones. Further, it depends neither on the discount factor, hence δ is not identified, nor on individual income realizations, hence φ and γ_y^2 are not identified, nor on the income processes, hence F_Y is irrelevant. However, it depends on $x_{t-1} = \lambda$, $\forall t$, which is not observable. That is, the co-state variable is constant and equal to the initial relative Pareto weight. This means that $x_{it} = \lambda_i$, $\forall t$, in equation (11).

---

9.   Below, θ often denotes a subset of these parameters, and is used as a short form for "parameters to be estimated".

Taking the logarithm of the optimality condition (11), with $x_{it} = \lambda_i$, with respect to (true) consumption and rearranging, one obtains

$$\sigma_i \log c_{it} = \sigma \log c_{vt} + \log \lambda_i.$$

Next, one can eliminate the unobservable term $\log \lambda_i$ by taking first differences. This gives

$$\sigma_i (\log c_{it} - \log c_{i,t-1}) = \sigma (\log c_{vt} - \log c_{v,t-1}). \tag{12}$$

It is clear that, given the consumption growth of household v ("the village"), the consumption growth of household i depends only on the ratio σ_i / σ, and not on σ_i and σ separately. This means that a normalization is needed. I normalize the average coefficient of RRA in each village to 1, that is, $\sigma = 1$.[10] That is, $u_v() = \log()$ and $\sigma_i = 1 + z_i' \beta + \varepsilon_i^\sigma$. Each β_k is identified from the correlation between the corresponding covariate, z_k, and the variability of consumption across households.[11]

In terms of measured consumptions today, c_{it}^* and $c_{vt}^* = (\sum_i c_{it}^*)/N$, (12) is

$$\left(1 + z_i' \beta + \varepsilon_i^\sigma\right) \log c_{it}^* - \log c_{vt}^* = \left(1 + z_i' \beta + \varepsilon_i^\sigma\right) \log c_{i,t-1} - \log c_{v,t-1}$$
$$+ \left(1 + z_i' \beta + \varepsilon_i^\sigma\right) \varepsilon_{it}^c - \varepsilon_{vt}^c. \tag{13}$$

Note that (13) is similar to a first-differenced equation often used to test perfect risk sharing, controlling for village–time effects and household characteristics. Tests in the vein of Townsend (1994) would include a measure of idiosyncratic risk in that equation and test whether its coefficient is zero.

Let $\varepsilon_{t-1}^{c,j}$ denote an N-vector of realizations of measurement error in households' consumption at time $t-1$, where each element $\varepsilon_{i,t-1}^{c,j}$ is drawn from $N(0, \gamma_c^2)$. Let $\varepsilon^{\sigma,j}$ denote an N-vector of realizations of the unobservable part of the coefficient of RRA. I first write the likelihood of each it observation conditional on $(\varepsilon_{t-1}^{c,j}, \varepsilon^{\sigma,j})$. Then, averaging the conditional likelihood over J draws, I integrate them out. That is, I use a simulated (pseudo) maximum likelihood estimator (SMLE).[12]

Let

$$\left(\psi_i^{prs,j}\right)^2 (Z, \theta) \equiv \left[\left((N-1)/N + z_i' \beta + \varepsilon_i^{\sigma,j}\right)^2 + (N-1)/N^2\right]\gamma_c^2$$

---

10.   This is true even with a long panel; see Chiappori et al. (2013), who have chosen the same normalization.

11.   Suppose, for example, that the consumption of households with high z_k, for some k, is smoother, i.e., they bear less consumption risk. According to the model, this is because they are more risk averse, hence β_k is positive. Chiappori et al. (2013) estimate a σ_i for each household using an 84-wave panel from the Townsend Thai Monthly Survey. This is not possible with the shorter panel of the ICRISAT data set used in this paper, see what follows.

12.   Note that it is not necessary to use simulation to take into account measurement error in consumption at time $t-1$ in the perfect risk-sharing case. I do it to be consistent with the LC case.

and

$$d_{it}^{prs,j}(Z,\theta) \equiv \left[\left(1 + z_i'\beta + \varepsilon_i^{\sigma,j}\right) \log \left(\frac{c_{it}^*}{c_{i,t-1}^* / \exp\left(\varepsilon_{i,t-1}^{c,j}\right)} \right) \right.$$
$$\left. - \log \left(\frac{c_{vt}^*}{c_{v,t-1}^* / \exp\left(\varepsilon_{v,t-1}^{c,j}\right)} \right) \right] \bigg/ \psi_i^{prs,j}(Z,\theta).$$

Then, we may write the (pseudo) likelihood of observation *it* conditional on $(\varepsilon_{t-1}^{c,j}, \varepsilon^{\sigma,j})$ as $\varphi(d_{it}^{prs,j}(Z,\theta))$, where φ is the density of the standard normal distribution. Finally, making J draws, the SMLE maximizes

$$\ell^{prs}(Z,\theta) = \sum_{i=1}^{N} \sum_{t=2}^{T} \log \left[\frac{1}{J} \sum_{j=1}^{J} \varphi\big(d_{it}^{prs,j}(Z,\theta)\big) \right] \tag{14}$$

with respect to θ, that is, β and γ_c^2. I also estimate the model without preference heterogeneity for comparison. This means setting $\beta = 0$ and $\gamma_\sigma^2 = 0$.

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

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

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).

3.2. Risk Sharing with Limited Commitment

Recall that in the LC case, the (true) ratio of marginal utilities from the previous period, x_{t-1}, is a sufficient statistic for everything that happened in the past. Hence, instead of conditioning on the history of shocks, s^t, and λ, it is sufficient to condition on the current income state, s_t, and x_{t-1}. This means that $x_{it} = x_{it}(s_t, x_{t-1})$ in equation (11).

Recall that $\hat{x}_{it}$ and $\hat{c}_{it}$ denote the predicted relative Pareto weight and consumption of household i at time t. Note that these are functions of (true) income, x_{t-1}, household characteristics, and parameters, that is, of $(y_t, x_{t-1}, Z; \theta, \gamma_\sigma^2, F_Y)$. Further, in general, they cannot be expressed analytically, therefore it is necessary to use numerical methods.

Next, I show that the structural parameters are identified, under some conditions. Then, Section 3.2.2 details how I estimate the model, including how I deal with measurement error.

3.2.1. Identification. When risk sharing is perfect, the predicted consumption allocation is independent of average risk aversion in the village, σ, the discount factor, δ, and the punishment parameter, φ. The question is whether one can identify these parameters in the case of risk sharing with LC.[13] I assume that some but not perfect risk sharing occurs in the data. For simplicity, I consider two ex-ante identical households. The heterogeneity parameters, β, are identified in the case of perfect risk sharing as well.

CLAIM 1. *Assume that at least one PC is binding for each household in the long run. Then, given two parameters among δ, σ, and φ, the third is identified.*

Proof. In Online Appendix B.   □

CLAIM 2. *Assume three possible income realizations and that at least two PCs are binding for each household in the long run. Then, given δ, σ, or φ, the other two parameters are jointly identified.*

Proof. In Online Appendix B.   □

CONJECTURE 1. *Assume that at least three PCs are binding for each household in the long run. Then, δ, σ, and φ are jointly identified.*

In the case where three PCs are binding, I provide a three-equation system which pins down the three parameters, see Online Appendix B. However, I cannot show uniqueness analytically. Numerically the three parameters are uniquely determined by the system.

3.2.2. Estimation. Replacing for x_{it} in (11) gives

$$\frac{(c_{vt})^{-\sigma}}{(c_{it})^{-\sigma_i}} = \hat{x}_{it}(y_t, x_{t-1}, Z; \theta, F_Y) = \frac{\hat{c}_{vt,i}(y_t, x_{t-1}, Z; \theta, F_Y)^{-\sigma}}{\hat{c}_{it}(y_t, x_{t-1}, Z; \theta, F_Y)^{-\sigma_i}}, \quad \forall i.$$

In terms of measured consumptions today, taking logarithms, replacing $\sigma + z_i'\beta + \varepsilon_i^\sigma$ for σ_i, dividing by σ, and rearranging, we have

$$\left(1 + \frac{z_i'\beta + \varepsilon_i^\sigma}{\sigma}\right) \log c_{it}^* - \log c_{vt}^*$$

$$= \left(1 + \frac{z_i'\beta + \varepsilon_i^\sigma}{\sigma}\right) \log \hat{c}_{it}(y_t, x_{t-1}, Z; \theta, F_Y) - \log \hat{c}_{vt,i}(y_t, x_{t-1}, Z; \theta, F_Y)$$

$$+ \left(1 + \frac{z_i'\beta + \varepsilon_i^\sigma}{\sigma}\right) \varepsilon_{it}^c - \varepsilon_{vt}^c. \tag{15}$$

---

13.   The income processes, F_Y, which are irrelevant in the perfect risk-sharing case but needed now, are estimated using the income data only and standard methods (see Online Appendix C).

Further, current incomes are measured with error too, and instead of $x_{i,t-1}$ only

$$x_{i,t-1}^* = \frac{\left(\exp(\varepsilon_{v,t-1}^c)\right)^{-\sigma}}{\left(\exp(\varepsilon_{i,t-1}^c)\right)^{-(\sigma+z_i'\beta+\varepsilon_i^\sigma)}} x_{i,t-1} \tag{16}$$

is "observed". That is, measurement errors in both income and consumption and unobserved heterogeneity in the coefficient of RRA influence the updating of the state variable. Note that, contrary to tests of perfect risk sharing in the vein of Townsend (1994), the model with LC implies a nonlinear relationship between current consumption shares and past consumption shares, as well as between current consumption shares and current income shares.

Let $(\varepsilon_t^{y,j}, \varepsilon_{t-1}^{c,j})$ denote an $N \times 2$ matrix of realizations of measurement errors in households' income at time t and consumption at time $t-1$, where each element $\varepsilon_{it}^{y,j}$ is drawn from the distribution $N(0, \gamma_y^2)$, and each element $\varepsilon_{i,t-1}^{c,j}$ is drawn from $N(0, \gamma_c^2)$. Remember that $\varepsilon^{\sigma,j}$ denotes an N-vector of realizations of the unobservable part of the coefficient of RRA. Knowing $x_{i,t-1}^*$, $\varepsilon_{i,t-1}^{c,j}$, and $\varepsilon_i^{\sigma,j}$, and using $\varepsilon_{t-1}^{c,j}$ to find $\varepsilon_{v,t-1}^{c,j}$, one can use equation (16) to compute $x_{i,t-1}$.[14] Similarly, knowing y_{it}^* and $\varepsilon_{it}^{y,j}$, one can easily compute y_{it}.

Let

$$(\psi_i^{lc,j})^2(Z,\theta) \equiv \left[\left((N-1)/N + \frac{z_i'\beta + \varepsilon_i^{\sigma,j}}{\sigma}\right)^2 + (N-1)/N^2\right]\gamma_c^2$$

and

$$
\begin{aligned}
d_{it}^{lc,j} &\left(y_t^*, x_{t-1}^*, Z; \theta, F_Y\right) \\
&\equiv \left[\left(1 + \frac{z_i'\beta + \varepsilon_i^{\sigma,j}}{\sigma}\right) \log\left(\frac{c_{it}^*}{\hat{c}_{it}\left(y_t^*, \varepsilon_t^{y,j}, x_{t-1}^*, \varepsilon_{t-1}^{c,j}, Z; \theta, F_Y\right)}\right)\right. \\
&\left. \quad - \log\left(\frac{c_{vt}^*}{\hat{c}_{vt,i}\left(y_t^*, \varepsilon_t^{y,j}, x_{t-1}^*, \varepsilon_{t-1}^{c,j}, Z; \theta, F_Y\right)}\right)\right] \Big/ \psi_i^{lc,j}(Z,\theta). \quad (17)
\end{aligned}
$$

---

14. With measurement error, all past values of consumption could be informative of x_{t-1}. For tractability, I only deal with the density (10).

Then, the (pseudo) likelihood of observation it conditional on $(\varepsilon_t^{y,j}, \varepsilon_{t-1}^{c,j}, \varepsilon^{\sigma,j})$ is $\varphi(d_{it}^{lc,j}(y_t^*, x_{t-1}^*, Z; \theta, F_Y))$.[15] Making J draws, the SMLE maximizes

$$\ell^{lc}(\theta) = \sum_{i=1}^{N} \sum_{t=2}^{T} \log \left[\frac{1}{J} \sum_{j=1}^{J} \varphi \left(d_{it}^{lc,j} \left(y_t^* \, x_{t-1}^*, Z; \theta, F_Y \right) \right) \right]. \tag{18}$$

Allowing for misspecification, the SMLE consistently estimates the pseudo-true values of the parameters and is asymptotically normal as both the number of it observations, denoted M, and the number of simulations, J, tend to infinity, and $\sqrt{M}/J \to 0$ (see Gouriéroux and Monfort 1997, for example).

The estimation involves three steps.

1. Estimation of the income processes. Online Appendix C gives the details.

2. Inner optimization. This step involves solving the model of risk sharing with LC to find the state-dependent optimal intervals, given σ_i, σ, δ, φ, F_{Y_i}, and F_{Y_v}, $\forall i$. Online Appendix A presents the details. Then, given the state-dependent optimal intervals, observed consumptions at time $t-1$ and incomes at time t, and draws for measurement errors, the predicted consumption allocation can be computed as described at the end of Section 2.2.[16]

3. Outer optimization. The log-likelihood (18) is maximized over the structural parameters, $\theta = (\sigma, \beta, \delta, \varphi, \gamma_c^2, \gamma_y^2)$. For comparison, the model is also estimated without preference heterogeneity, that is, imposing $\beta = 0$ and $\gamma_\sigma^2 = 0$.

Iterating between the dynamic program which solves the model and the likelihood maximization routine is computationally very costly. Instead, I solve the model on a four-dimensional grid for σ_i, σ, δ, φ, given F_{Y_i} and F_{Y_v}. The optimal intervals given any parameter values are then computed by linear interpolation. As a result, the model does not have to be solved again and again while maximizing equation (18) with respect to θ. To find the parameter estimates, I iterate between a standard optimization algorithm available in R[17] and grid search in order to find the global optimum.

The estimation of the model of risk sharing with LC involves both simulation and approximation. As Ackerberg, Geweke, and Hahn (2009) point out, in terms of the asymptotic properties of the maximum likelihood estimator, approximation error in computed dynamic models has similar effects as a limited number of simulations,

---

15.   Note that, in the consumption insurance literature, preference shocks are often used to account for the error term in the estimating equation, or, as in Cochrane (1991), consumption growth is measured with error. These alternative assumptions are not suitable in the case of risk sharing with LC. The former because today's shock would drop out of the estimating equation (15), while in the latter case one would have to draw the previous period's measurement error, $\varepsilon_{t-1}^{c,j}$, from a random walk.

16.   The number of simulations, J, is 50, hence $J > \sqrt{M}$, because I have maximum 185 it ($t = 2, \ldots, T$) observations per village.

17.   Namely, L-BFGS-B, a quasi-Newton method which allows box constraints on parameters.

and the results from the literature on simulated maximum likelihood estimation apply
(Hajivassiliou and Ruud 1994; Gouriéroux and Monfort 1997).

When computing the variance–covariance matrix, the information matrix equality
is not assumed to hold, and possible serial correlation is taken into account, as in the
perfect risk-sharing case.

3.3. Model Selection

To statistically compare the models, I use a model selection test introduced by Vuong
(1989). Vuong proposes a general test to compare two (nested or nonnested) models
to determine which model is closer to the true data-generating process, allowing for
the possibility that neither model is correctly specified.

The test is based on the difference between the log likelihood values of the two
models. Suppose that we want to compare model 1 and model 2 using M observations.
Denote the log likelihood of observation m for model 1 (resp. 2) at the estimated
parameter vector by ℓ_m^1 (resp. ℓ_m^2). The likelihood ratio is defined as

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

Denote the number of parameters to be estimated by q_1 (resp. q_2) for model 1
(resp. 2).

Under the null that the two models are equally close to the data,

$$2LR \ \Rightarrow \ {}_w\chi^2_{q_1+q_2}(\cdot; \hat{\kappa}),$$

where $\Rightarrow$ means convergence in distribution and ${}_w\chi^2_{q_1+q_2}(\cdot; \hat{\kappa})$ is the cumulative
distribution function of the weighted χ^2 distribution—that is, a weighted sum the
squares of $(q_1 + q_2)$ standard normal variables. The weights $\hat{\kappa}$ are 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

$$\hat{A}^1 = \sum_{i=1}^{N} \sum_{t=2}^{T} -\nabla_\theta^2 \ell_{it}^1,$$
$$\hat{B}^1 = \sum_{i=1}^{N} \sum_{t=2}^{T} \hat{s}_{it}^1 \hat{s}_{it}^{1'},$$

where $\hat{s}_{it}^1$ is the estimated score for household i at time t for model 1, and similarly
for model 2, and

$$\hat{B}^{1,2} = \hat{B}^{2,1'} = \sum_{i=1}^{N} \sum_{t=2}^{T} \hat{s}_{it}^1 \hat{s}_{it}^{2'}.$$

The p-values of the weighted χ^2 distribution have to be simulated. I do 100,000 replications.[18]

I compare four models: risk sharing with LC with heterogeneous preferences (LC^{u_i}) and with homogeneous preferences (LC^u), and perfect risk sharing with heterogeneous preferences (PRS^{u_i}) and with homogeneous preferences (PRS^u). The heterogeneous versions of both models nest their homogeneous versions, and LC models nest their perfect risk-sharing counterparts with $\varphi \rightarrow 1$.[19]

4. Application

This section applies the estimation and model selection methods discussed previously to study risk sharing in three Indian villages.

4.1. Data

I use data from the Village Level Studies conducted by the International Crop Research Institute for the Semi-Arid Tropics (ICRISAT) in India from 1975 to 1984.[20] It is safe to say that it is the most widely used income–consumption survey from developing countries. In particular, the ICRISAT data set has been used by many papers studying risk sharing in village economies, including Townsend (1994), Ligon (1998), Ogaki and Zhang (2001), LTW, and Mazzocco and Saini (2012). I focus on three villages, Aurepalle, Kanzara, and Shirapur, and the years 1976 to 1981, because of concern over the accuracy of measured consumption in the other years (Townsend 1994; LTW).[21]

---

18.   Vuong (1989) proposes a more powerful test for nonnested models. In particular, $LR/\sqrt{M}\hat{\omega} \Rightarrow N(0, 1)$, where $\hat{\omega}$ is the estimated standard deviation of the likelihood ratio, that is,

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

Given that most pairwise model comparisons in this paper involve nested models, I focus on the general test.

19.   Without the possibility for large punishments for deviating, the LC models do not nest their perfect risk-sharing counterparts, because even as the discount factor approaches 1, perfect risk sharing is self-enforcing only if the relative Pareto weight is in a certain interval. In other words, the LC model (with ex-post PCs) nests a model of risk sharing with ex-ante PCs, but not the perfect risk-sharing model with any Pareto weights.

20.   I thank ICRISAT for making the data available, Reena Badiani and Ethan Ligon for making their constructed aggregates available, and Maurizio Mazzocco and Shiv Saini for sharing data construction codes.

21.   A new data set with a bigger sample spanning more than 10 years (six of which are publicly available) with frequent interviews has been collected, namely the Townsend Thai Monthly Survey. However, household demographics, such as age and gender of household members, are not publicly available for each period. See Attanasio (1999), for example, on the importance of accounting for household demographics when studying consumption.

TABLE 1. Descriptive statistics.

Variable	Aurepalle		Kanzara		Shirapur	
	Mean	SD	Mean	SD	Mean	SD
Aeq. nondur. cons.[a]	303.47	127.86	400.84	161.42	430.37	170.71
Aeq. income[a]	629.58	429.78	984.42	742.54	792.16	577.57
Education[b]	2.27	1.66	3.41	1.61	3.67	1.31
Proportion of women	0.49	0.26	0.54	0.23	0.50	0.25
Age	53.43	11.45	43.51	9.97	49.60	9.58
Aeq. land[c]	0.62	0.65	0.83	0.93	1.14	1.22
# of observations	204		222		186	
# of households	34		37		31	

a. Measured in 1975 Indian rupees per year. In 1975, approximately 8 Indian rupees were worth 1 US dollar, which is about 4 dollars in 2013.
b. Measured on a scale $1 =$ illiterate to $8 =$ more than a first degree.
c. Measured in hectares.

The nondurable consumption measure I use includes food consumption, clothing, services, utilities, and narcotics. Income includes net income from crop production, labor, livestock, and transfers from outside the village. Both consumption and income used in the analysis are yearly and per adult equivalent. To compute the adult-equivalent size of each household, I use the same age–gender weights as Townsend (1994).[22] Preference heterogeneity is captured by four variables: education, gender, age, and land as a proxy for initial wealth. That is, I include these variables as z_i. Education is the longitudinal average of the highest education level in the household. Gender is measured by the average proportion of women among adults.[23] Age is the age of the head of the household at the initial period.[24] To capture the effect of initial wealth, I include the size of land operated at the initial period.[25] I do not aim to find the best way to capture differences in the curvature of the current utility function across households, due to its computational burden, only to see whether allowing for preference heterogeneity improves the models' fit to the data in a statistical and an economic sense, and whether predicted policy effects differ.

Table 1 presents descriptive statistics for the three villages. On average, daily nondurable consumption per adult equivalent is 0.83, 1.10, and 1.18 1975 Indian rupees in Aurepalle, Kandara, and Shirapur, respectively, which is about 0.42, 0.55, and

22.   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.

23.   I include this last variable instead of a female head dummy, because there are very few households headed by a woman in rural India.

24.   Chiappori et al. (2013) find, using the Townsend Thai Monthly Survey, that age is most correlated with households' risk aversion assuming that risk sharing is perfect.

25.   LTW speculate that allowing for increasing relative risk aversion may help the model capture both the dynamic response of consumption to idiosyncratic income shocks and cross-sectional consumption inequality.

0.59 2013 US dollars, respectively. The difference between nondurable consumption reflects durable consumption and investment.

4.2. Reduced-Form Results

Before turning to the structural results, I present reduced-form evidence on how nondurable consumption reacts to idiosyncratic income shocks. I regress the first difference of the logarithm of adult-equivalent consumption on the first difference of the logarithm of adult-equivalent income, controlling for year dummies. I find that a 1% increase in household income per adult equivalent leads to an increase of 0.206, 0.222, and 0.169% (with robust and clustered standard errors of 0.056, 0.065, and 0.057, which imply p-values of 0.000, 0.001, and 0.000) in household nondurable consumption per adult equivalent in Aurepalle, Kandara, and Shirapur, respectively, controlling for aggregate consumption.[26] These parameter estimates suggest that risk sharing is partial rather than perfect. However, households in these three villages achieve a remarkable amount of insurance: only about one-fifth of idiosyncratic income fluctuations translate into consumption fluctuations. Note that I have not included any time-varying observables in these regressions. They serve to describe the dynamic response of consumption to idiosyncratic income shocks. In what follows I examine whether LC models are able to capture this response of consumption.

4.3. Structural Estimation and Model Selection Results

As a first step, households' income processes have to be estimated, see Section 3.2.2.[27] In the baseline specification, I estimate the risk-sharing models with measurement error in income and last period's consumption using an SMLE, but without an unobservable term in the curvature of the utility function.[28]

Table 2 shows the baseline structural estimation and the model selection results for all three villages. The bottom panel shows the model selection tests, conducting Vuong's general test (see Section 3.3) to statistically compare the models for the three villages. Three conclusions can be drawn from the results. (i) Models with LC explain the consumption data significantly better compared to their perfect risk-sharing counterparts. That is, we can reject perfect risk sharing against a fully specified model of partial mutual insurance. (ii) Models with preference heterogeneity are able to fit the consumption data significantly better than their homogeneous counterparts. (iii) The PRS^{u_i} model outperforms the LC^u model in a statistical sense.

---

26.   Note that these estimated coefficients are higher than the ones found by Townsend (1994). This is due to differences in the measurement of income and consumption (see also Ravallion and Chaudhuri 1997).

27.   Online Appendix C shows the estimated parameters of the income processes. It also explains how the processes are discretized.

28.   As robustness checks I also estimate the models (i) without measurement error in income and last period's consumption, and (ii) with measurement errors, as in the baseline, and adding an unobservable term in the curvature of the utility function with $\gamma_\sigma = 0.02$. These robustness checks are presented in Online Appendix D.

TABLE 2. Structural estimation and model selection results with measurement errors..

	LC^u_i	LC^u	PRS^u_i	PRS^u	LC^u_i	LC^u	PRS^u_i	PRS^u	LC^u_i	LC^u	PRS^u_i	PRS^u
δ (discount factor)	0.937***	0.980***			0.939***	0.791***			0.917***	0.908***		
	(0.004)	(0.042)			(0.000)	(0.001)			(0.000)	(0.001)		
φ (punishment parameter)	0.371***	0.338***			0.098***	0.256***			0.262***	0.062***		
	(0.001)	(0.024)			(0.000)	(0.001)			(0.007)	(0.001)		
σ (average coef. of RRA)	1.084***	3.387***			0.941***	1.448***			1.180***	3.206***		
	(0.000)	(0.071)			(0.000)	(0.000)			(0.009)	(0.001)		
β_1/σ (education)	0.040***		0.439***		0.163***		0.119***		0.504***		0.219***	
	(0.004)		(0.052)		(0.001)		(0.041)		(0.014)		(0.081)	
β_2/σ (proportion of women)	1.374***		−0.075		−1.600***		−0.777***		3.240***		1.739***	
	(0.003)		(0.145)		(0.004)		(0.315)		(0.017)		(0.307)	
β_3/σ (age)	−0.037***		−0.003		0.023***		0.039***		−0.044***		−0.029***	
	(0.001)		(0.003)		(0.000)		(0.006)		(0.001)		(0.004)	
β_4/σ (land)	−0.179***		−0.027		−0.427***		−0.137***		−0.630***		0.015	
	(0.002)		(0.088)		(0.001)		(0.027)		(0.009)		(0.030)	
γ_c^2 (variance fo cons. m.e.)	0.043***	0.036***	0.049***	0.037***	0.043***	0.044***	0.043***	0.044***	0.053***	0.048***	0.052***	0.049***
	(0.001)	(0.009)	(0.008)	(0.005)	(0.001)	(0.001)	(0.010)	(0.016)	(0.007)	(0.001)	(0.008)	(0.008)
γ_y^2 (variance fo inc. m.e.)	0.058***	0.057***			0.050***	0.089***			0.086***	0.121***		
	(0.000)	(0.001)			(0.000)	(0.000)			(0.000)	(0.000)		
Log likelihood	14.414	−6.450	1.245	−15.822	12.168	−7.116	−6.214	−32.202	31.013	−31.826	−9.957	−37.367
# of obs.	170	170	170	170	185	185	185	185	155	155	155	155
# of hholds	34	34	34	34	37	37	37	37	31	31	31	31

Vuong's test – $H_{all.}$: model of column is closer to the true data generating process than model of row

	LC^u_i	LC^u	PRS^u_i	PRS^u	LC^u_i	LC^u	PRS^u_i	PRS^u	LC^u_i	LC^u	PRS^u_i	PRS^u
LC^u	41.729***		15.391***		38.569***		1.804***		125.678***		43.737***	
	(0.000)		(0.006)		(0.000)		(0.000)		(0.000)		(0.000)	
PRS^u_i	26.338***				36.735***				81.941***			
	(0.000)				(0.001)				(0.000)			
PRS^u	60.474***	18.745***	34.135***		88.741***	50.172**	51.976***		136.761***	11.084***	54.820***	
	(0.000)	(0.000)	(0.000)		(0.006)	(0.043)	(0.001)		(0.000)	(0.006)	(0.000)	

Notes: LC^u_i: equation (15) is estimated by maximizing the log-likelihood function (18). LC^u: equation (15) with $\beta = 0$ is estimated. PRS^u_i: equation (13) is estimated by maximizing the log-likelihood function (14). PRS^u: equation (13) with $\beta = 0$ is estimated. The β for the LC model are divided by σ to make them comparable with the estimates for the perfect risk-sharing model. 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, where the alternative hypothesis is that the model of the column is closer to the true data generating process than the model of the row. The p-values are simulated using 100,000 replications.
***Significant at 1%.

The top panel of Table 2 shows the structural parameter estimates. Allowing for preference heterogeneity leads to remarkably stable estimates for both δ and σ across the three villages: the former varies between 0.917 and 0.939, and the latter between 0.941 and 1.180. Hence, these data suggest that the coefficient of RRA is close to 1, and that households in these villages discount the future at about 8%, a rate twice as high as what is usually assumed. The estimated δ and σ vary greatly across the three villages when preferences are assumed homogeneous, between 1.448 and 3.387 and 0.791 and 0.980, respectively. The estimates for φ, the fraction of consumption lost in autarky, vary substantially across villages for both LC models, hence we learn little about this parameter other than that it is not zero.

Turning to the heterogeneity parameters, according to the estimates in Table 2, households' coefficient of RRA depends positively on education and negatively on land operated in all three villages (except for the perfect risk model in Aurepalle and Shirapur where land is not significant), and positively on the proportion of women and negatively on age in Aurepalle and Shirapur (except for the perfect risk model in Aurepalle where neither is significant), while the opposite is true in Kanzara.[29,30]

Next, I evaluate the models in terms of how well they can capture consumption dynamics and the cross-sectional variation of consumption. First, I check whether changes in income explain the difference between changes in true consumption and changes in predicted consumptions. If income is not significant in such a regression, then the corresponding model has been able to account for the dynamic response of consumption to idiosyncratic income shocks. Table 3 shows the results. For the LC^{u_i} model changes in income are not significant at the 5% level. The point estimates for the coefficients on income changes are about half of those in Section 4.2. The perfect risk-sharing models, not surprisingly, fail this test, and the LC^u model cannot account for the dynamic response of consumption to income in two of the three villages.

Second, I compute what fraction of the variation of consumption unexplained by household and time effects each model captures. In order to do this, I regress the residuals from regressions of predicted consumption from each model on household and time dummies on the residuals from regressing true consumption on household and time dummies. The R^2 of these regressions tell us how well each model does according to this criterion. Table 4 presents the numbers. For the first two villages, the LC^{u_i} model is able to capture about 21.7% of the variation of consumption unexplained by household and time effects. The other models also do well for Aurepalle. All models do badly according to this measure for Shirapur.

---

29.   Unfortunately, it is not easy to determine which features of the data cause these two parameters to switch sign for Kanzara. The heterogeneity parameters capture that (i) more risk-averse households should bear less uninsurable risk, and (ii) because of LC less risk-averse households should have higher average consumption for given consumption risk and bear less consumption risk given average consumption, because the value of (risky) autarky consumption is higher for them.

30.   Online Appendix E presents descriptive statistics for the coefficients of RRA from the two models with heterogeneous preferences to get a sense of how much preference heterogeneity the estimated parameters imply.

TABLE 3. Do the models capture the dynamic response of consumption to idiosyncratic income shocks?

Dependent variable: Δ log true cons. $-\Delta$ log predicted cons. Explanatory variable: Δ log income

	Aurepalle				Kanzara				Shirapur			
	LC^{u}_{i}	LC^{u}	PRS^{u}_{i}	PRS^{u}	LC^{u}_{i}	LC^{u}	PRS^{u}_{i}	PRS^{u}	LC^{u}_{i}	LC^{u}	PRS^{u}_{i}	PRS^{u}
	0.091*	0.192***	0.227***	0.206***	0.106	0.202***	0.284***	0.222***	0.077	0.080*	0.144***	0.169***
	(0.062)	(0.056)	(0.058)	(0.056)	(0.064)	(0.063)	(0.060)	(0.064)	(0.063)	(0.046)	(0.045)	(0.047)

Notes: Robust and clustered standard errors are in parentheses. Time dummies are included in all regressions.
***Significant at 1%; *significant at 10%.

TABLE 4. What fraction of the variation of consumption can the models explain above what is possible with household and time dummies?

Aurepalle				Kanzara				Shirapur			
$LC^u{}_i$	LC^u	$PRS^u{}_i$	PRS^u	$LC^u{}_i$	LC^u	$PRS^u{}_i$	PRS^u	$LC^u{}_i$	LC^u	$PRS^u{}_i$	PRS^u
0.213	0.218	0.157	0.202	0.220	0.080	0.094	0.076	0.000	0.001	0.015	0.000

Note: R^2 from regressing the residuals from a regression of predicted consumption on household and time dummies on the residuals from a regression of true consumption on household and time dummies.

TABLE 5. How well can the models account for consumption inequality and changes in inequality?

	Variance of log consumption					Gini coefficient				
	Data	$LC^u{}_i$	LC^u	$PRS^u{}_i$	PRS^u	Data	$LC^u{}_i$	LC^u	$PRS^u{}_i$	PRS^u
Average inequality	0.17	0.19	0.13	0.17	0.15	0.21	0.21	0.19	0.21	0.20
Corr(true inequality, predicted inequality)		0.82	0.44	0.60	0.63		0.79	0.13	0.19	0.30
Corr(Δ true inequality, Δ predicted inequality)		0.60	−0.07	0.15	0.31		0.07	−0.17	−0.04	−0.21

Third, I compute two measures of cross-sectional consumption inequality for the data and consumption predicted by each model—namely, the variance of log consumption and the Gini coefficient (see, e.g., Krueger et al. 2010). I then compute (i) average inequality, (ii) the correlation between true inequality and inequality predicted by each model, and (iii) the correlation between changes in true inequality and changes in inequality predicted by each model (i.e., predicted inequality minus true inequality in the previous period). Since this results in only five observations for each village, Table 5 presents the results combining the three villages.

When the variance of log consumption is used as a measure of inequality, the LC^u model underpredicts consumption inequality on average (as in Krueger and Perri 2006), while the heterogeneous version overpredicts it. The perfect risk-sharing models do better in capturing inequality on average. This is not surprising, however, since with many observations they would exactly match average inequality. When the Gini coefficient is used as a measure of inequality, the prediction for average inequality from the $LC^u{}_i$ model almost exactly matches that of the data. The correlation between predicted and true inequality and the correlation between predicted and true changes in inequality are in favor of the $LC^u{}_i$ model. The former correlation is 82.4% and 79.0% for the variance of log consumption and the Gini coefficient, respectively, while the latter is as high as 59.8% when the variance of log consumption measures inequality. The LC^u model completely fails at predicting changes in inequality.

TABLE 6. The effects of a counterfactual tax and transfer policy.

Average change in...		Predicted cons.		Predicted CE cons.		Consumption (direct redist.)	Income
		LC^{u_i}	LC^u	LC^{u_i}	LC^u		
Aurepalle	Poor	2.19	−3.46	2.96	−1.81	10.66	32.01
	Rich	−2.19	3.46	−4.15	9.89	−10.66	−32.01
Kanzara	Poor	6.08	1.95	6.72	2.45	15.69	45.08
	Rich	−6.42	−2.05	−5.94	−1.90	−16.56	−47.58
Shirapur	Poor	8.21	8.16	7.86	8.65	21.26	34.10
	Rich	−8.76	−8.71	−7.87	−7.58	−22.64	−36.37

Notes: Poor households are those with average income below the village median. CE consumption = certainty-equivalent consumption.

4.4. Policy Simulation

This section uses the estimated structural models to examine the effects of a counterfactual tax and transfer policy. I consider income taxation at a 30% rate, the revenues of which are given out as lump-sum transfers—that is, the policy is progressive and revenue-neutral for each village. Note that the policy can be thought of as public insurance (Krueger and Perri 2011; Broer 2011). I study the effects of this policy on the consumption of both poor and rich households. The poor are defined as households with average income below the village median.[31] I simulate the effects of the policy considering each possible time of introduction, $t = 2, \ldots, 6$, assuming in each case that the policy will remain in place in all future periods, and average over the predicted consumption changes at the time of introduction.

According to perfect risk-sharing models, redistributive policies have no effect on the consumption allocation. This is because aggregate consumption does not change, and consumption shares are given by predetermined Pareto weights. On the contrary, LC models are able to predict a redistribution of consumption as a result of income redistribution. Table 6 summarizes the simulation results for the LC models, both in terms of changes in predicted consumption (columns 1 and 2) and changes in certainty-equivalent (CE) consumption (columns 3 and 4), a measure of welfare. For comparison, Table 6 also presents the effects of the policy on consumption under the assumption that the tax and transfer policy can be applied to consumption directly (column 5) and on income (column 6).

The LC models predict that the increase in consumption by the poor is about 20% to 40% of the increase of direct consumption redistribution. LC models are able to predict an increase in consumption by the poor, because the value of their outside option increases more than that of the rich as a result of the tax and transfer policy. Further, allowing for preference heterogeneity matters quantitatively: ignoring it would result

31. Median income is 237.4, 322.4, and 356.4 rupees per year per adult equivalent in Aurepalle, Kanzara, and Shirapur, respectively. Poor households' average income is 138.4, 242.9, and 321.0, respectively. Average income of the rich is 351.8, 551.7, and 556.0, respectively.

in predicting lower benefits for the poor by about 4.5 rupees in certainty-equivalent consumption per adult equivalent when conducting ex-ante policy evaluation for two of the three villages. This is because the poor are more risk averse given the estimated heterogeneity parameters (see Online Appendix E), which means that the reduction in income risk implied by the policy increases their autarky utility more.[32]

5. Concluding Remarks

This paper has compared four models of risk sharing in terms of how well they can explain the consumption allocation in three Indian villages. The results suggest that both limitations to the enforcement of informal risk-sharing contracts and heterogeneity in risk preferences are important in explaining consumption shares in a statistical sense in all villages studied. The limited commitment model with heterogeneous preferences also outperforms the benchmark models in terms of (i) explaining the response of consumption to idiosyncratic income shocks, (ii) accounting for the variation of consumption unexplained by household fixed effects and time dummies, and (iii) capturing the variation of inequality across time and villages and predicting changes in inequality, as measured by the variance of log consumption and the Gini coefficient.

Using the structural estimation results, this paper has then simulated the effects on the consumption allocation of a counterfactual tax and transfer policy. Perfect risk-sharing models predict no change in the consumption allocation as a result of such a policy, while models with limited commitment are able to predict an increase in consumption by the poor. Allowing for preference heterogeneity is quantitatively important when predicting the effects of the policy. Research on the structural modeling of how consumption is allocated among households in poor communities can serve as an input for policy evaluation and design. Policy-makers and members of nongovernmental organizations could have a better understanding of the effects of their programs, such as redistributive policies or micro-insurance programs, by taking into account existing informal arrangements to share risk.

Several interesting extensions are possible. First, whether heterogeneity in the discount factor across households is important should also be addressed. Second, other models of mutual risk sharing could be incorporated into the analysis, such as a model of risk sharing with private information (Wang 1995).[33] Another important task for future work is to bring models with limited commitment and individual saving (Ligon, Thomas and Worrall 2000; Ábrahám and Laczó 2013) to the data. Finally, when complete markets do not exist to insure against income fluctuations, households are expected to smooth their income, not just consumption (Morduch 1995). This could be

---

32.   However, this only matters if participation constraints of poor households bind sufficiently often relative to those of the rich. For Shirapur the predicted consumption changes are almost identical with and without preference heterogeneity. This is because in this village the participation constraints of some rich households with low risk aversion bind relatively more often.

33.   In a recent paper, Kinnan (2010) finds that asymmetric information about income realizations is important in accounting for partial insurance in Thai villages. Whether such a model is useful for predicting the consumption allocation and quantitative policy effects is to be studied.

formalized in the context of this paper, endogenizing income by allowing households to choose between several income generating processes. Then, the cost of imperfect consumption insurance in terms of lower expected incomes could be quantified.

References

Ábrahám, Árpád and Sarolta Laczó (2013). "Efficient Risk Sharing with Limited Commitment and Storage." Barcelona GSE Working Paper No. 697.

Ackerberg, Daniel, John Geweke, and Jinyong Hahn (2009). "Comments on "Convergence Properties of the Likelihood of Computed Dynamic Models" by Fernández-Villaverde, Rubio-Ramírez and Santos." *Econometrica*, 77, 2009–2017.

Altug, Sumru and Robert A. Miller (1990). "Household Choices in Equilibrium." *Econometrica*, 58, 543–570.

Attanasio, Orazio and José-Víctor Ríos-Rull (2000). "Consumption Smoothing in Island Economies: Can Public Insurance Reduce Welfare?" *European Economic Review*, 44, 1225–1258.

Attanasio, Orazio P. (1999). "Consumption." In *Handbook of Macroeconomics* (Vol. 1), edited by J.B. Taylor and M. Woodford. Elsevier, pp. 741–812.

Blundell, Richard, Luigi Pistaferri, and Ian Preston (2008). "Consumption Inequality and Partial Insurance." *American Economic Review*, 98(5), 1887–1921.

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

Broer, Tobias (2011). "Crowding Out and Crowding In: When Does Redistribution Improve Risk-Sharing in Limited Commitment Economies?" *Journal of Economic Theory*, 146, 957–975.

Chiappori, Pierre-André, Krislert Samphantharak, Sam Schulhofer-Wohl, and Robert M. Townsend (2013). "Heterogeneity and Risk Sharing in Village Economies." Federal Reserve Bank of Minneapolis Research Department Staff Report No. 483.

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

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

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

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

Dubois, Pierre, Bruno Jullien, and Thierry Magnac (2008). "Formal and Informal Risk Sharing in LDCs: Theory and Empirical Evidence." *Econometrica*, 76, 679–725.

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

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

Gouriéroux, Christian and Alain Monfort (1997). *Simulation-Based Econometric Methods*. Oxford University Press.

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

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

Hajivassiliou, Vassilis A. and Paul A. Ruud (1994). "Classical Estimation Methods for LDV Models Using Simulation." In *Handbook of Econometrics*, Vol. 4, edited by Robert F. Engle and Daniel L. McFadden. Elsevier, pp. 2383–2441.

Karaivanov, Alexander (2012). "Financial Constraints and Occupational Choice in Thai Villages." *Journal of Development Economics*, 97, 201–220.

Karaivanov, Alexander and Robert M. Townsend (2014). "Dynamic Financial Constraints: Distinguishing Mechanism Design from Exogenously Incomplete Regimes." *Econometrica*, 82, 887–959.

Kehoe, Patrick J. and Fabrizio Perri (2002). "International Business Cycles with Endogenous Incomplete Markets." *Econometrica*, 70, 907–928.

Kinnan, Cynthia (2010). "Distinguishing Barriers to Insurance in Thai Villages." CFSP Working Paper.

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

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

Krueger, Dirk and Fabrizio Perri (2011). "Public versus Private Risk Sharing." *Journal of Economic Theory*, 146, 920–956.

Krueger, Dirk, Fabrizio Perri, Luigi Pistaferri, and Giovanni L. Violante (2010). "Cross-sectional Facts for Macroeconomists. *Review of Economic Dynamics*, 13, 1–14.

Krusell, Per and Anthony A. Smith (1998). "Income and Wealth Heterogeneity in the Macroeconomy." *Journal of Political Economy*, 106, 867–896.

Ligon, Ethan (1998). "Risk Sharing and Information in Village Economies." *Review of Economic Studies*, 65, 847–864.

Ligon, Ethan, Jonathan P. Thomas, and Timothy Worrall (2000). "Mutual Insurance, Individual Savings, and Limited Commitment." *Review of Economic Dynamics*, 3, 216–246.

Ligon, Ethan, Jonathan P. Thomas, and Timothy Worrall (2002). "Informal Insurance Arrangements with Limited Commitment: Theory and Evidence from Village Economies." *Review of Economic Studies*, 69, 209–244.

Mace, Barbara J. (1991). "Full Insurance in the Presence of Aggregate Uncertainty." *Journal of Political Economy*, 99, 928–956.

Marcet, Albert and Ramon Marimon (2011). "Recursive Contracts." CEP Discussion Paper No. 1055.

Mazzocco, Maurizio and Shiv Saini (2012). "Testing Efficient Risk Sharing with Heterogeneous Risk Preferences." *American Economic Review*, 102(1), 428–468.

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

Ogaki, Masao and Qiang Zhang (2001). "Decreasing Relative Risk Aversion and Tests of Risk Sharing." *Econometrica*, 69, 515–526.

Paulson, Anna L., Robert M. Townsend, and Alexander K. Karaivanov (2006). "Distinguishing Limited Liability from Moral Hazard in a Model of Entrepreneurship." *Journal of Political Economy*, 114, 100–144.

Ravallion, Martin and Shubham Chaudhuri (1997). "Risk and Insurance in Village India: Comment." *Econometrica*, 65, 171–184.

Schulhofer-Wohl, Sam (2011). "Heterogeneity and Tests of Risk Sharing." *Journal of Political Economy*, 119, 925–958.

Todd, Petra E. and Kenneth I. Wolpin (2006). "Assessing the Impact of a School Subsidy Program in Mexico: Using a Social Experiment to Validate a Dynamic Behavioral Model of Child Schooling and Fertility." *American Economic Review*, 96(5), 1384–1417.

Todd, Petra E. and Kenneth I. Wolpin (2008). "Ex Ante Evaluation of Social Programs." *Annals of Economics and Statistics*, 91/92, 263–291.

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

Udry, Christopher (1994). "Risk and Insurance in a Rural Credit Market: An Empirical Investigation in Northern Nigeria." *Review of Economic Studies*, 61, 495–526.

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

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

Wilson, Robert (1968). "The Theory of Syndicates." *Econometrica*, 36, 119–132.
Wooldridge, Jeffrey M. (2002). *Econometric Analysis of Cross Section and Panel Data*. MIT Press.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Online Appendix
Codes