Risk Sharing in Village Economies Revisited

Tessa Bold, IIES, Stockholm University and CEPR
Tobias Broer, Paris School of Economics, IIES, Stockholm University and CEPR*

September 2021

Abstract

We quantitatively evaluate a model of insurance with limited commitment where the requirement that contracts be immune to deviations by subcoalitions makes group size endogenous, as proposed by Genicot and Ray (2003). We compare the model's predictions to panel data from rural Indian villages. Apart from predicting a realistic degree of insurance, the model captures the evidence along two new dimensions: first, the largest coalition-proof groups are substantially smaller than typical villages. Second, with strong insurance in small groups, individual consumption responds symmetrically to income rises and falls, while alternative models predict strong counterfactual asymmetry.

JEL Classification: D11, D12, D52
Keywords: Risk Sharing, Endogenous Groups, Village Economies, Informal Insurance, Dynamic Limited Commitment, Renegotiation-Proofness

*We thank Árpád Ábrahám, Paul Klein, Per Krusell, Nicola Pavoni, Swapnil Singh, Christian Stoltenberg, Kjetil Storesletten and Nicholas Barton, as well as the editor and five referees, and seminar participants at Bocconi University, IIES Stockholm University, Normac, the SED annual meeting 2015, and the University of Oslo for comments. Contact: tessa.bold@iies.su.se, tobias.broer@iies.su.se.

1. Introduction

We structurally estimate a model of insurance in village economies with endogenous risk sharing groups as proposed by Genicot and Ray (2003). Apart from the observed degree of insurance, the model predicts, first, equilibrium groups that are substantially smaller than typical villages; and second, a symmetric reaction of consumption to positive and negative income shocks. This is not trivial, because other popular models of risk sharing predict counterfactual asymmetries in consumption-income comovement, and are silent about group size.

The key friction that enables us to study endogenous groups is the absence of commitment to co-insurance. This is a plausible reason for limited risk sharing in poor villages, where contract enforcement is difficult but other impediments to insurance, such as lack of information on households' productive possibilities and effort, are less pronounced. Moreover, previous studies have shown how limits to commitment can explain the partial risk sharing observed in many agricultural villages (Townsend, 1994; Ligon et al., 2002; Laczó, 2014). To study endogenous group formation in a fully dynamic and quantitative model of risk sharing with limited commitment to contracts, we assume that households can renege on village insurance not just alone but in subgroups, as in Genicot and Ray (2003). We think the resulting requirement of 'coalition-proofness' is particularly appealing in the context of village economies, where it seems difficult to prevent those who renege on insurance arrangements to insure each other again in the future.

Our first contribution is to draw attention to group size as an important determinant of risk sharing, and to propose a tractable way to make insurance groups endogenous outcomes of a dynamic limited commitment risk sharing mechanism. To compute the risk-sharing equilibrium quantitatively, we combine the common approximative solution of the standard limited commitment model, originally proposed by Ligon et al. (2002) and used, for example, in Laczó (2014) and Dubois et al. (2008), with the recursive procedure for finding stable group sizes when coalitions can deviate together proposed by Genicot and Ray (2003).

We use this to estimate the model for the well-known ICRISAT dataset on agricultural villages in India.[1]

Our second contribution is to show how the model with endogenous groups replicates the degree of risk sharing in those villages well and captures the empirical evidence along two new dimensions: first, it predicts insurance groups whose size is in line with the single-digit groups documented in other datasets, and with evidence on bilateral consumption correlations in the ICRISAT data (which, however, lacks explicit information on groups).[2] And second, it captures the approximate symmetry of empirical consumption-income comovements. This is important because the limited commitment constraint per se is more likely to bind for villagers with high income realisations and therefore attractive outside options. In large insurance groups, such as countries, this feature is known to imply a much stronger response of consumption to positive than to negative

1. Using a simplified, stationary version of the model, Dubois (2006) and Fitzsimons et al. (2015) test empirically whether risk sharing at the village level is constrained by coalitional deviations. Both papers find evidence for coalitional deviations, but neither estimates the conditional consumption distribution that would arise if risk sharing was restricted by them. Bold (2009) derives a formal test of the presence of coalitional deviations in a dynamic setting that relies on the finding that groups that are constrained endogenously do not exhibit the amnesia typical of the standard model (Kocherlakota, 1996), but requires an exact identification of constrained households.

2. The standard punishment assumption of eternal individual autarky implies that there is no limit to the size of insurance groups, as a larger group size increases the benefits of co-insurance but does not affect the outside option. This partly motivates the common focus on 'village-level' risk sharing. Transfers are typically made, however, in groups that are much smaller than the village, giving rise to a small recent literature that focuses on limited commitment to bilateral insurance relationships. The focus on the resulting network structure, however, makes the analysis of truly dynamic constrained risk sharing infeasible, requiring instead a static and, usually, exogenous risk sharing rule (Bloch et al., 2008; Ambrus et al., 2014). Importantly, our finding of small groups does not contradict the finding of insurance within (larger) kinship groups, castes etc. (Angelucci et al., 2015; Fitzsimons et al., 2015; Mazzocco and Saini, 2012; Mobarak and Rosenzweig, 2012), to the extent that these may contain several smaller endogenous groups that feature strong risk sharing and, for reasons we do not consider here, may not typically cross caste or kinship barriers.

income shocks (as the former make the outside option more attractive and thus tighten participation constraints, while the latter do not) that is not seen in the data (Broer, 2013). Beyond pairs, where consumption shares trivially move in symmetry (Kocherlakota, 1996), however, the strength of this asymmetry both in theory and data has so far been unknown for small communities. We show that in the standard version of the model, without coalitional deviations, the asymmetry implied by the limited commitment constraint increases quickly with the exogenous group size. And at the typical village sizes considered in the literature, it is in fact substantially larger than in the ICRISAT data. In contrast, the model with coalitional deviations predicts negligible asymmetry. This is substantially closer to the ICRISAT data, where the asymmetry is negative but for the most part small and insignificant.

To show that these results do not depend on the particular environment we consider, we show that they are robust to alternative assumptions about outside options, the income process, and to including stylised preference heterogeneity. Moreover, we show that the standard computational approximation of the equilibrium that we use (Ligon et al., 2002) produces results very similar to those from an exact computation. The largest groups for which we can compute the risk-sharing contract exactly is, with 4 households, equal to the sizes we estimate using the approximation. In addition, we show that the largest renegotiation-proof groups are essentially the same as in our benchmark results (where insurance is strong but not perfect) when computing them exactly under the maintained assumption of full insurance (where we can consider any group size).

The next section introduces the dynamic limited commitment model with coalitional deviations, and describes our quantitative approximation. Section 3 describes the ICRISAT data and estimates the strength of insurance and its asymmetry. Section 4 presents the estimation results for the model with coalitional deviations, and compares them to those for the standard limited commitment model with individual deviations, and a model of self insurance. Section 5 provides additional evidence for small insurance groups in the ICRISAT data. It also compares the predictions of the coalitional deviations model to those of the standard limited commitment model when group sizes are small in both models. An online appendix contains robustness checks and additional analysis.

2. Consumption insurance with limited commitment

This section describes a dynamic model of co-insurance under limited commitment with endogenous group formation. We also present two alternatives, a model with exogenous group formation and, in Section 2.6, a model where agents self-insure via savings.

The setting for mutual insurance is an economy where risk-averse households face idiosyncratic income risk but cannot commit to making the transfers implied by risk sharing. Consumption insurance is thus restricted by (ex-post) participation constraints: the utility value of continued participation in an insurance scheme must not be less than that of households' outside option in any state of the world.

2.1. A limited commitment village economy

We consider a village community with N households. In each period $t = 1, 2, ...\infty$, household i receives an endowment y_t^i of the only consumption good. These endowments are independent realizations of a discrete Markov chain that takes n different values $y_1 < y_2 < ... < y_n$, is identical across households and has transition probability from income value j to income value k given by p_{jk}. We denote by s_t the $N - vector$ of income realizations in the village in period t, which we henceforth call "state of nature". The Markov chain for individual incomes induces a (n^N-state) Markov chain for s_t, with π_{sr} the probability of transition from state s to state r. Households are infinitely lived and discount the future with a common discount factor δ. They have identical and twice continuously differentiable utility functions $u(\cdot)$ defined over consumption $c^i(s_t)$ in state s_t. Households are risk-averse and would therefore find it profitable to enter into a risk sharing arrangement with other villagers in order to smooth consumption in the face of idiosyncratic income movements. Households have perfect information about both their own income realizations and those of other villagers, but are not able to write binding contracts.

These assumptions of identical preferences and income processes across households imply that households are ex-ante identical and differ ex-post only in their income histories. The environment is thus less general than in Ligon et al. (2002), who allow for heterogeneous income processes and preferences. When comparing the model to the data in Section 4, we deal with unmodelled heterogeneity in the latter by purging it of both aggregate time variation and time-invariant individual heterogeneity in consumption and income.

Households can enter an insurance arrangement with $n \leq N$ households. An insurance arrangement consists of a list of net transfers $\tau^1, ..., \tau^n$ in each period. These transfers are feasible if $\sum_{i=1}^{n} \tau^i = 0$ and $\tau^i \leq y^i \; \forall i$. We denote by h_t the history of states up to period $t-1$ and by H_t the set of all possible histories up to date t (h_0 is simply the empty set). An insurance contract, σ, is a sequence of functions, $\{\sigma_t\}_{t=0}^{\infty}$ that maps $H_t \times \mathcal{S}$ to a list of feasible transfers. We can then define the expected utility from an insurance contract σ after history h_t when the current state of the world is s as

$$U_s^i(h_t) = u(c_s^i(h_t)) + E \sum_{j=t+1}^{\infty} \delta^{j-t} u(c_r^i(h_j)) \tag{1}$$

where $c_s^i(h_t) = y^i(s_t) - \tau_s^i(h_t)$ and the expectations operator E denotes the expectation with respect to the distribution over the product of possible histories and current states r in period j.

The limited commitment environment we consider does not allow for ex-ante binding agreements to be written. Instead, in each period and after the realization of the current state, households decide whether to conform with the insurance arrangement (and make the specified transfer) or to renege on their transfer. In such an environment, insurance contracts must be self-enforcing. This implies that following any history h_t and current realization of the state, household i's expected discounted utility from conforming with the insurance contract must not be smaller than some possibly state-dependent utility outside the contract V_s^i. That is, a contract is considered sustainable if for all households i

$$U_s^i(h_t) \geq V_s^i \quad \forall h_t, s \in \mathcal{S} \quad . \tag{2}$$

Insurance transfers from households with high income realizations to those with low income realizations can be sustained in such a context whenever non-participation in the contract is costly, implying that the instantaneous benefit from deviating from the contract is traded off against a lower continuation value under the outside option V_s^i, for example, because it implies the loss of future insurance possibilities.

To find the constrained-optimal insurance contract for a general outside option V_s^i, we can write down the dynamic programme that solves for the Pareto frontier in an insurance group of size n. In particular, we maximise the utility of agent n taking as state variables the promised life-time utilities of the other $n-1$ agents, which summarise the history of states and transfers up to the current period (Abreu et al., 1990; Ligon et al., 2002).

The constrained-optimal contract is the solution to the following Lagrangian:

$$U_s^n(U_s^1, U_s^2, ..., U_s^{n-1}) = \max_{((U_r^i)_{r=1}^S)_{i=1}^{n-1}, (c_s^i)_{i=1}^n} u(c_s^n) + \delta \sum_{r=1}^S \pi_{sr} U_r^n \left(U_r^1, ..., U_r^{n-1}\right) \tag{3}$$

subject to a set of promise-keeping constraints

$$\gamma^i : \qquad u(c_s^i) + \delta \sum_{r=1}^S \pi_{sr} U_r^i \geq U_s^i \quad \forall i \neq n, \tag{4}$$

a set of enforcement constraints

$$\delta\gamma^i \pi_{sr} \varphi_r^i : \qquad U_r^i \geq V_r^i \quad \forall i, r \in \mathcal{S} \tag{5}$$

and an aggregate resource constraint in each state and period.

$$\omega : \qquad \sum_{i=1}^n y_s^i \geq \sum_{i=1}^n c_s^i \tag{6}$$

where γ^i, φ^i, and ω are the Lagrange multipliers associated with the promise-keeping, enforcement, and resource constraints respectively.

The first-order and envelope conditions associated with this problem imply the following optimality condition that links the evolution of household i's consumption between state s in period t and state r in period $t+1$ to that of a reference household n.

$$\gamma_r^i = \frac{u'(c_r^n)}{u'(c_r^i)} = \frac{1+\varphi_r^i}{1+\varphi_r^n}\gamma^i = \frac{1+\varphi_r^i}{1+\varphi_r^n}\frac{u'(c_s^n)}{u'(c_s^i)} \quad \forall r \in \mathcal{S}, \quad \forall i \neq n. \tag{7}$$

Equation (7) captures the essence of consumption insurance with limited commitment to contracts: relative marginal utility is constant between state s in period t and state r in period $t+1$, and insurance thus perfect, as long as participation constraints are not binding. Now suppose agent i has a binding constraint, $\varphi_r^i > 0$. From equation (7), this implies the relative marginal utility in state r tomorrow is higher than the relative marginal utility in state s today. By the concavity of the utility function, this implies that agent i increases their consumption relative to agent n. Thus, the model generates partial insurance with both perfect insurance and autarky as two limiting cases.

Inherent in the model is an asymmetry for the consumption process that is most easily illustrated with log-preferences, where the relative consumption of any two households c_r^i/c_r^j equals their relative, 'updated' Lagrange multipliers $\gamma^i(1+\varphi_r^i)/\gamma^j(1+\varphi_r^j)$. Summing across all households i in period t and using the resource constraint $Y_r = \sum_{i=1}^n c_r^i$ we can express household j's consumption as a function of village income Y_r and Lagrange multipliers: $c_r^j = Y_r \frac{\gamma^j(1+\varphi_r^j)}{\sum_{i=1}^n \gamma^i(1+\varphi_r^i)}$. Taking log-differences of both sides yields

$$d\log(c_t^j) = d\log(Y_t) + \log(1+\varphi_t^j) - \log\left(1 + \frac{\sum_{i=1}^n \gamma^i \varphi_t^i}{\sum_{i=1}^n \gamma^i}\right). \tag{8}$$

where $d\log$ denotes the log difference and we suppress the dependence on state r in period t. Individual consumption growth is thus the sum of three terms: first, it is proportional to output growth $d\log(Y_t)$; second, it has an individual-specific term $\log(1+\varphi_t^j) \geq 0$ that is positive when agent j has a binding constraint and the multiplier φ_t^j is positive, but zero otherwise; and finally, there is a 'drift-term' $-\log(1+\frac{\sum_{i=1}^n \gamma^i\varphi_t^i}{\sum_{i=1}^n \gamma^i}) \leq 0$ that is common for all group members and strictly negative whenever at least one participation constraint is binding in the group.

Equation (8) illustrates the asymmetry inherent in risk sharing in a limited commitment environment: the consumption share of household i increases only when its participation constraint binds. Moreover, for a given vector of outside options of other villagers, its consumption share is increasing in her outside option V_r^i. Whenever the participation constraint is slack, the household shares the same decline in marginal utility with other unconstrained households, where the magnitude of the decline is independent of its outside option.

Apart from the assumption of log-preferences and i.i.d. transitions, these analytical results require that we can distinguish constrained from unconstrained individuals. Importantly, however, for many variants of the limited commitment model, the right-hand side of the participation constraint (5) is increasing in the individual endowment y_s^i. It is thus more likely to bind after income increases and this has motivated researchers to look for asymmetries in the joint distribution of consumption and income.

Figure 1 illustrates this asymmetry in a limited commitment model similar to that analyzed in Ligon et al. (2002), but simplified by our assumptions of identical income processes and preferences (in contrast to their setup with heterogeneous income processes and preferences). Ligon et al. (2002) assume that the insurance group coincides with the village and the outside option to the contract is eternal individual autarky. The expected utility of this outside option thus equals

$$V_s^i = V(y_s^i) = u(y_s^i) + \delta \sum_{r=1}^{S} \pi_{sr} V_r^i \ \forall i, s \tag{9}$$

and therefore, for typical income processes, increases in current income realizations y_s^i will increase the outside option, meaning that binding constraints will coincide with periods of rising income. In what follows, we refer to this model as the 'Individual Deviation' or 'ID' model.[3] Its implied joint distribution of consumption and income growth, depicted as a scatter plot in Figure 1, shows a pronounced kink in both the conditional mean and variance functions around zero income growth, consistent with the intuition suggested by equation (8) and in line with previous studies of economies with larger populations (Krueger and Perri, 2005; Broer, 2013).

2.2. Consumption-risk sharing in endogenous groups

We now present a model, first introduced by Genicot and Ray (2003), of coalition-proof dynamic risk sharing. The environment is the same as in Section 2.1: risk-averse and patient agents face a volatile income stream, which they can smooth by entering into a mutual insurance arrangement that must be self-enforcing. Just as in the ID model discussed in Section 2.1, an agent's outside option always

3. Using the approximation in Ligon et al. (2002), we solve a simplified version of the model that abstracts from heterogeneity in income processes and direct utility penalties after a deviation, with risk aversion equal to 0.95 and a discount factor equal to 0.85, equal to the authors' estimates when targeting the observed log-changes in consumption and incomes in the ICRISAT household panel data set, and an income process estimated on data for the ICRISAT village of Aurepalle after partialling out time fixed effects and controlling for household fixed effects (see Section 3 for more details on the data).

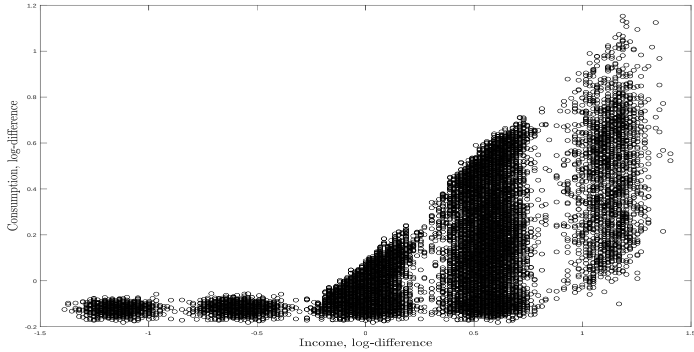

FIGURE 1. Consumption and income growth in general equilibrium

Note: The figure shows a scatter plot of consumption and income growth (or their log-differences) for a simplified version of the model presented in Ligon et al. (2002), and their parameter estimates corresponding to the village of Aurepalle in the ICRISAT data set. Specifically, we use preferences with constant relative risk aversion equal to 0.95 and a discount factor of 0.85, equal to their estimates when targeting the observed log-changes in consumption and incomes. We use a simplified version of the model that abstracts from heterogeneity in income processes and direct utility penalties after a deviation. The figure plots residuals from a regression that controls for movements in aggregate resources.

involves less opportunity for consumption smoothing than the existing contract under consideration. But the outside option does not equate to eternal autarky when insurance groups are larger than two households. Instead, the idea is that any subgroup of agents can insure each other forever after. We thus replace the outside option of eternal autarky with continued insurance, albeit in a smaller group.[4] We call this model the 'CD' model (for 'coalitional deviations').

We now show how to assess stability of an insurance group of size n relative to such group deviations. To do so, we must define outside options for groups of $m < n$ players. Beginning with an individual, their expected life-time utility

4. More formal treatments of such coalition-proof equilibria are given in Farrell and Maskin (1989); Bernheim et al. (1987); Bernheim and Ray (1989).

starting in state s is given by

$$V_s^i(1) = u(y_s^i) + \delta \sum_{r=1}^{S} \pi_{sr} V_r^i(1) \tag{10}$$

where relative to equation (9) we have introduced a dependence on the size of the deviating group. Just as in the individual deviations model, this payoff defines the outside option for an individual who considers deviation from a group of size n.

The point of departure in the CD model comes when considering deviations by groups of size $m > 1$. In order to define outside options in the case of coalitional deviations for a subgroup of size m, we must first define the payoffs this group can achieve when sharing risk. To this end, we proceed both recursively and by example to arrive at the general definition of stability given by Genicot and Ray (2003).

For $m = 2$, consider a history-dependent contract σ that maps the history of the game h_t and current state into a set of feasible transfers, $\tau_s^i(h_t)$ for $i = 1, 2$. We can then write the payoff for the two players from an insurance scheme σ starting after history h_t and current state s as

$$U_s^i(h_t, \sigma, 2) = u(c_s^i) + \delta \sum_{r=1}^{S} \pi_{sr} U_r^i(h_{t+1}, \sigma, 2) \tag{11}$$

for $i = 1, 2$. Stability of an insurance scheme is then assessed according to the following criterion: there must not be a history h_t, current state s, and associated payoffs prescribed by the contract σ following this history and state, such that for either $i = 1, 2$

$$V_s^i(1) > U_s^i(h_t, \sigma, 2) \tag{12}$$

If the contract σ satisfies this criterion, then we say that it is stable. Moreover, we let $U_s^i(h_0, \sigma, 2)$ be the expected life-time payoff assigned to agent i in this contract when the current state is s and h_0 is the history at date 0, which is simply the empty set. We denote by $\mathbf{V}_s(\sigma, 2)$ the 2×1 vector that contains these payoffs for the two agents. Let the set $\mathcal{V}_s^*(2)$ contain all such stable payoff vectors for a group of size 2.

Consider then a group of size 3 and the stability of contracts in this group, where again a contract is a sequence of maps from the history of the game h_t

into a set of feasible transfers, $\tau_s^i(h_t)$ for $i = 1, 2, 3$. A contract for a group of size 3 is then considered stable if there is no history h_t, current state of the world s, and associated payoffs prescribed by the contract σ following this history and state such that a subgroup has a profitable deviation. That is, it must not be the case that for any $i = 1, 2, 3$

$$V_s^i(1) > U_s^i(h_t, \sigma, 3), \tag{13}$$

nor must it be the case that there is a subgroup of size $m = 2$ and $\mathbf{V}_s \in \mathcal{V}_s^*(2)$, such that

$$V_s^i > U_s^i(h_t, \sigma, 3) \tag{14}$$

for all i in the subgroup. If the contract σ satisfies these criteria, then we say that it is stable. Moreover, we let $U_s^i(h_0, \sigma, 3)$ be the expected life-time payoff assigned to agent i in this contract when the current state is s and h_0 is the history at date 0 (the empty set). We denote by $\mathbf{V}_s(\sigma, 3)$ the 3×1 vector that contains these payoffs for the three agents and let the set $\mathcal{V}_s^*(3)$ contain all such stable payoff vectors for a group of size 3. If there is no contract that satisfies the above criteria, then these sets are empty.

Having defined sets of stable payoff vectors for all groups of size $m = 1, ..., n-1$, and hence their outside options, we can now give the general definition of stability for an insurance group of size n. That is, a contract σ is stable if there is no history h_t, current state s and continuation payoffs prescribed by the contract following this history and state, such that for some subgroup of individuals (of size $m < n$) and some stable expected payoff vector $\mathbf{V}_s \in \mathcal{V}_s^*(m)$

$$V_s^i > U_s^i(h_t, \sigma, n) \;\; \forall i \text{ in the subgroup.} \tag{15}$$

Again, if the contract σ satisfies the criterion, it is considered stable. As before, we denote by $\mathbf{V}_s(\sigma, n)$ the $n \times 1$ vector that contains the expected life-time utilities of the n agents when the initial state is s and collect all such stable payoff vectors in the set $\mathcal{V}_s^*(n)$. If there is no contract for a group of size n that is stable, then the group is unstable.

The planner's problem then boils down to finding the best contract (that generates a particular stable payoff vector) among the stable ones in a group of size n by maximising the expected life-time utility of agent n subject to agent $1, ..., n-1$ receiving their promised utilities $U^1, ..., U^{n-1}$. That is, the planner solves the following dynamic optimization:

$$U_s^n(U_s^1, U_s^2, ..., U_s^{n-1}) = \max_{((U_r^i)_{r=1}^S)_{i=1}^{n-1}, (c_s^i)_{i=1}^n} u(c_s^n) + \delta \sum_{r=1}^S \pi_{sr} U_r^n \left(U_r^1, ..., U_r^{n-1}\right) \tag{16}$$

subject to a set of promise-keeping constraints

$$u(c_s^i) + \delta \sum_{r=1}^S \pi_{sr} U_r^i \geq U_s^i \quad \forall i \neq n, \tag{17}$$

the enforcement constraints

$$(U_r^i)_{i=1}^n \in \mathcal{V}_r^*(n) \quad \forall r \in \mathcal{S} \tag{18}$$

and an aggregate resource constraint in each state and period.

$$\sum_{i=1}^n y_s^i \geq \sum_{i=1}^n c_s^i \quad . \tag{19}$$

The coalition-proof outside options differ from the standard ID model in two ways: first, by making deviations – to a new insurance group rather than individual autarky – more attractive, they restrict risk sharing more strongly. Second, because the coalition-proof equilibrium risk-sharing contract may fail to exist for a given group size and the maximal sustainable size of an insurance group is known to be bounded (Genicot and Ray, 2003), the model endogenizes the size of insurance groups (or more precisely, the size of the largest stable group in a community of size N).[5]

This contrasts with the ID model that has no mechanism to bound group size. Since the marginal benefit of adding members to the risk-sharing group is always positive (Murgai et al., 2002), this has led researchers to consider the largest

5. The intuition for the boundedness of groups in the model with coalitional deviations is complex, but relies on the fact that the insurance benefit of increasing group size goes to zero as groups become large. If the sustainable group size was infinite, one could choose a finite group such that it contains with probability 1 in every state a large enough coalition of n high income individuals that exhausts all diversification benefits. This coalition would have no incentives to make transfers, so autarky would be the only equilibrium in the larger group. See Genicot and Ray (2003), p. 94.

possible group that a household could interact with as the relevant insurance group. In a developing country context, this is often the village or smaller exogenously bounded groups within it such as extended families. In contrast, the CD model allows us to derive sharp predictions for the equilibrium group size.

2.3. Approximating the CD model

To test whether the CD model is a good quantitative model for village risk sharing, we need to find the set of stable insurance group sizes and contracts in a typical village. However, since the state space over promised continuation values (which encode the history of the contract) and possible income realizations grows exponentially with insurance group size n, even the standard ID model has so far been solved only for groups of up to three agents. This has led researchers to consider approximate solutions to the constrained-efficient contract.

Relative to the ID model, the CD model introduces an additional layer of complexity: in contrast to the individual deviations model, which gives each individual one threat point (conditional on the current state), a deviating subgroup can now continue to share risk starting from any division of surplusses in its set of stable expected payoff vectors. Since this set typically has infinitely many members, this leads to an infinite number of potential threats and possible strategies for the planner to deter them. In either model, an exact numerical solution for larger group sizes is out of reach.

To operationalize the CD model for quantitative analysis, we therefore adapt and extend the common approximation to the solution of the standard ID version. This approximation, originally proposed by Ligon et al. (2002) and used, for example, in Laczó (2014) and Dubois et al. (2008), reduces the dimensionality of finding the constrained-efficient risk-sharing allocation in a village of N members by recasting the simultaneous $N-$ household insurance problem as a sequence of N two-player problems, in which an individual shares risk with an agent who represents the rest of the village of $N-1$ individuals, and who has average preferences and receives an endowment equal to the average across the $N-1$ remaining villagers. The vector of outside options for the individual and the representative rest-of-village agent equal the consumption value of individual and average incomes respectively.

This approximation can be thought of as the constrained-efficient equilibrium of a simplified infinitely repeated insurance game where the planner abstracts from heterogeneity in the rest of the village when calculating the Lagrange multipliers φ_r^i in equation (5). Specifically, φ_r^i is derived in a simplified game where all other villagers $j \neq i$ are assigned a common multiplier, φ_r^{-i}, equivalent to pooling their income both inside the contract and during deviation. With the multipliers φ_r^i obtained in the simplified game in hand, the planner then solves the $N-1$ first-order conditions in equation (7) for consumption (thus never using the approximate multipliers φ_r^{-i}).

To adapt this standard approximation to the case of coalitional deviations, we combine the 'one-against-the-rest-of-village' strategy with a recursive identification of stable coalitions. Our aim is to define an outside option of an individual i sharing risk in a group of size n that captures the idea of coalition-proofness and the dynamic nature of the contract. At the same time, we are looking for an approximation that simplifies the dynamic contract along three dimensions: (1) the approximate solution procedure will not track the entire history of shocks and transfers of the group; (2) it will not consider the entire stable set $\mathcal{V}_s^*(m)$ following state s as potential deviations for a group of size m, (3) it will not consider all possible strategies for the planner to deter group deviations.

To this end, we consider insurance contracts in groups of n villagers that increase in size $n = 2, 3, ..., N$. As in the ID model, the shadow value φ_r^i of a given outside option of individual i in state r is found by solving n sequential two-agent games where individual $i = 1, ..., n$ interacts with a rest of the group that pools income on and off the equilibrium path. For given n, the outside option of the rest of the group is thus unchanged relative to the standard approximation (so the planner abstracts from heterogeneity within the rest of the group when determining the shadow value of its outside option, φ_r^{-i}).

For $n = 2$, the individual's outside option is autarky, as in the ID model. Importantly, for $n > 2$, there is a potentially infinite set of stable outside options $\mathcal{V}_s^*(m)$ for any group of size $1 < m < n$. To reduce the set of possible deviations for $n > 2$, we assume that members of a deviating coalition share utility in the most equal fashion as in Park et al. (2018). This corresponds to equal initial Lagrange multipliers associated with the promise-keeping constraint (17) for individual i and households in the rest of the group following a deviation (or equivalently in our environment equal Pareto weights for all deviators), subject to any binding participation constraints (that may arise immediately).

The "most equal sharing" assumption reduces the infinite set of stable outside options to a single vector. As in the ID model, outside options depend only on the vector of current incomes, but are independent of income and transfer histories and the household's identity. While in the ID model this independence follows entirely from the assumptions of identical income processes and preferences, it is, not without loss of generality in the CD model. In particular, an alternative choice of outside option would assume that deviators inherit the relative initial Lagrange multipliers from the n-household insurance arrangement (subject to binding participation constraints). While the history dependence this introduces to the outside option makes much of our analysis in this paper unfeasible, we do explore the robustness of our benchmark results to this alternative assumption in Section A.9.

The assumption that the rest of the village pools income perfectly reduces the information about the income distribution we need to keep track of. Specifically, it is only necessary to record the individual's income realization and the number, $j_l(s)$, (but not the identity) of individuals with income level y_l in the rest of the group when the state of the world is s. The state of the world is thus summarized by the vector $\{y_s^{ind}, j_1(s), ..., j_L(s)\}$.

With this notation in hand, we can write down the outside option for the rest of the group in an insurance group of size n when the current state of the world is s as:

$$V_{s,aut}^{rog}(n-1) = u(\bar{y}_s^{rog}) + \delta \sum_r \pi_{sr} V_{r,aut}^{rog}(n-1), \tag{20}$$

where $\bar{y}_s^{rog} = \frac{\sum_{l=1}^{L} j_l(s) y_l}{n-1}$ denotes the average income in the rest of the group.

Turning to the outside options of the individual in a contract $\hat{\sigma}$ in a group of size n, suppose that we have defined stable contracts $\hat{\sigma}$ for groups of size $m = 1, ..., n-1$ and associated sets of stable payoff vectors starting in state s, $\mathcal{V}_s^*(m)$ with typical element $\mathbf{V}_s = \{V_s^{ind}, V_s^{rog}\}$ consisting of the expected lifetime utility of the individual and the rest of the group. Within this set of stable payoffs, we denote the 'most equal' vector by $\bar{\mathbf{V}}_s = \{\bar{V}_s^{ind}, \bar{V}_s^{rog}\}$.

We can then define stability of an approximate contract in a group of size n. Such a contract is considered stable if neither the individual nor the rest-of-the-village have a profitable deviation following any history of transfers and states.

This requires (i) that the expected life-time utility of the rest-of-the group inside the contract is no smaller than deviation into autarky, that is, it must not be the case that:

$$U_s^{rog}(h_t, \hat{\sigma}, n) = u(c_s^{rog}) + \delta \sum_{r=1}^{S} \pi_{sr} U_r^{rog}(h_{t+1}, \hat{\sigma}, n) < V_{s,aut}^{rog}(n-1) \tag{21}$$

Further, for the individual to be willing to participate, there must not be a subset of size $m < n$ consisting of the individual and $m-1$ members of the rest of the group and payoff vector $\bar{\mathbf{V}}_s \in \mathcal{V}^*(m)$ such that both

$$U_s^{ind}(h_t, \hat{\sigma}, n) = u(c_s^{ind}) + \delta \sum_{r=1}^{S} \pi_{sr} U_r^{ind}(h_{t+1}, \sigma, n) < \bar{V}_s^{ind} \quad . \tag{22}$$

and

$$U_s^{rog}(h_t, \hat{\sigma}, n) < \bar{V}_s^{rog} \tag{23}$$

This definition deserves some further clarification. First, as already noted, the outside option of the rest of the group is the same as in the ID model and sustainable (or stable) contracts are structured such that the rest of the group is at least as well off under the contract as under their outside option. Thus, considerations of coalitional deviations only come into play when building the outside option of the individual. Specifically, stability requires that the individual is not worse off under the contract than what she could achieve by deviating together with a subset of the rest of the group and continuing insurance with the most equal allocation (condition (22)), provided that the subset of the rest of the group would want to participate in such a deviation (condition (23)). In other words, the individual cannot threaten to deviate with m members of the rest of the insurance group (and thus potentially improve her position inside the contract) unless this subset of the rest of the group would want to participate in such a deviation.

The final step solves for the constrained-efficient contract $\hat{\sigma}$ in the set of stable contracts between an individual and rest of the village of size n as n increases to encompass the village N. From these contracts, the planner finds the n consumption allocations for n individuals in a group of size n. The resulting consumption sequence will be non-stationary, in the sense that a history of transfers will – all else equal – be rewarded with a higher consumption allocation in the current period. The equilibrium contract will be symmetric, in the sense that agents with the same income realization and history of transfers (conditional on aggregate realizations) will be treated symmetrically.

2.4. Determination of group size

Once the set of stable group sizes and corresponding constrained-efficient contracts has been found, we need to determine how to distribute villagers in a village of size N into stable risk-sharing groups, which may be smaller than the village. We proceed as follows: when there are several stable group sizes, we denote the largest stable group by n^{max} and assume that k groups of size n^{max} are formed where k corresponds to the smallest integer such that $N' = k \times n^{max} \geq N$. That is, the allocation rule forms $k-1$ groups of size n^{max} from the existing villagers, forms the $k'th$ group from the remaining villagers, and, if necessary, increases village size by $N' - N \geq 0$ members to complete this last group.

We focus on the largest stable group for two reasons. First, this is the group size that implies the highest insurance benefit. So if individuals could choose group sizes ex ante, this is what they would choose. Second, this is consistent with the focus on the maximum group size in the ID model (where it is set equal to the village/sample size).

2.5. Discussion

It is important to stress that the model's focus is on group sustainability. The model is silent, however, on how sustainable groups are formed in the first place. While a formal treatment of this process is beyond the scope of this paper, we show in Bold and Broer (2018), in the context of an equal sharing insurance contract, conditions under which the set of stable sizes identified here coincides with the absorbing states of an equilibrium process of coalition formation (EPCF), as defined in Ray and Vohra (2015). As a more concrete example of how small risk-sharing groups can emerge in an environment of explicit group formation, Ambrus et al. (2014) show (albeit in a more restricted context than ours), how coalition-proof risk sharing in groups can emerge as the outcome of a decentralized model of bilateral transfers in a network.

It is also important to note that throughout the analysis, we maintain a number of simplifying assumptions. First, as in Genicot and Ray (2003), agents can form new subgroups only within an existing insurance group. We think this is reasonable as there may be many, unmodelled, reasons why (sub-)group-formation requires previous interaction. More importantly, group-formation without restriction typically causes problems for the formulation of a recursive solution concept and can lead to 'cyclical blocking chains'.[6] We therefore must for the moment rule out deviations with outsiders.

In our analysis of the CD and ID models we also abstract from accumulation of assets both at the individual and the group/village level. In a given insurance group, introducing the opportunity to save for individuals improves the outside option relative to financial autarky, particularly for the rich, and thus tightens participation constraints. The opportunity to accumulate assets at the village level, in contrast, improves insurance and thus relaxes constraints (Ligon et al., 2000; Ábrahám and Laczó, 2018). The net effect is thus ambiguous at given insurance group size, though in practice it has been shown that the introduction of savings increases consumption smoothing in a limited commitment environment (Chandrasekhar et al., 2010). With endogenous groups, in the CD model, all the same considerations apply for a given group size, but the effect would be even more complicated because maximum group sizes may increase or decrease with savings.

By changing the estimated discount factor, we would expect an ID model with savings to be able to produce the same amount of consumption smoothing as the benchmark estimation. Importantly, we do not think that the introduction of savings in the ID model will overturn our key result, which relies on the link between consumption smoothing and the asymmetry in the consumption response to income changes (see Figure 1), and which, as we will see, is also present in the self-insurance model in Section 2.6. The better empirical performance of the CD model, on the other hand, hinges on the substantially smaller size of insurance groups, and savings could worsen this if they substantially increase group size. There are, however, reasons to think that this will not be the case: Bold and Dercon (2014) show in simulation exercises of a stationary model of risk sharing under limited commitment with coalitional

6. See Genicot and Ray (2003), p. 97 for a discussion.

deviations, savings and state-independent penalties, that only small groups attain stability.

Migration of household members who expect to earn higher income in a nearby city, for example, adds additional insurance opportunities against common and idiosyncratic shocks and thus makes the risk-sharing scheme more attractive. It also makes the outside option of individual autarky more attractive, as households can use migration to insure against negative shocks. In the same context as ours, rural village India, Morten (2016) finds the net effect on risk sharing to be negative in a standard ID model with endogenous migration, but in the context of a large migration experiment in Bangladesh, it is positive (Meghir et al., 2019). Because the degree of insurance in the small groups we estimate in our CD model is higher, and participation constraints at low income (that migration tightens most strongly) thus bind less often, we might expect the negative effect to be less pronounced. Other related issues are how migration affects the sustainable group size, and whether permanent migration (which might be one reason for the attrition we observe in our data) would arise in equilibrium, and thus change group size over time.

We also assume that agents experience different income realisations but are otherwise identical, that insurance takes place in groups, and that deviating groups initially practice equal-sharing. We thus abstract from heterogeneity in income processes (Ligon et al., 2002) or preferences (Laczó, 2014), as well as limits to information within groups (Kinnan, 2014; Ligon, 1998). This is partly because additional dimensions of heterogeneity, frictions, and complexities would make the quantitative analysis of the model with coalitional deviations infeasible, but also because we believe that the effect of coalitional deviations is best highlighted in the most standard version of the limited commitment model.[7] In the Online Appendix, we relax a number of these assumptions and investigate how this affects the results.

7. For example, Laczó (2014) finds evidence of preference heterogeneity when estimating the standard limited commitment model with individual deviations. And Mazzocco and Saini (2012) reject the joint hypothesis of full insurance and homogeneous risk preferences for caste groups in the ICRISAT villages, but cannot reject full insurance when allowing for heterogeneity in risk preferences.

Our maintained assumptions also imply that we completely abstract from any network structure of the village or its subgroups. In fact, we view our work as complementary to studies analysing the formation of insurance networks with limited commitment (Bloch et al., 2008; Ambrus et al., 2014) where the focus on the structure of stable networks, however, requires a simplification of the analysis along dimensions that are central to our study.

2.6. Self-insurance

In the quantitative analysis, we will compare the CD model to two other models of consumption smoothing: first, the standard ID model, and second a simple self-insurance (henceforth SI) model where, instead of engaging in mutual insurance, households build a buffer against income shocks by accumulating savings $b_t^i \geq 0$ with a village lender, remunerated at an exogenous interest rate R. Their period budget constraint in period t is thus

$$c_t^i = y_t^i + Rb_{t-1}^i - b_t^i \tag{24}$$

We think of this model more as a useful, standard comparison, rather than one that captures the particular institutions in the ICRISAT villages that our quantitative analysis focuses on.

3. The data

This section introduces the village economies that have been used most widely to study models of risk sharing: the ICRISAT panel. We describe the data and show scatter plots and key moments, motivated from the theory presented in Section 2.1 and 2.2 that will allow us to choose a good quantitative model for village risk sharing. Since the ICRISAT data do not identify risk sharing units within the village, we, like many previous studies, focus on the joint distribution of individual consumption and income growth to evaluate risk sharing in the data. In contrast to previous work that concentrated on measures of the degree of insurance, we also study moments that capture the asymmetry suggested by the limited commitment mechanism in Section 2.1, namely the difference in comovement between consumption and income for those with income gains versus income losses. We confirm the strong degree of risk sharing found in previous studies of rural village economies. We also show that there are no positive asymmetries in income and consumption comovement.[8]

The data come from the village level studies conducted by the International Crop Research Institute for the Semi-Arid Tropics (ICRISAT) in India from 1975-1984. We focus on three rural and agricultural villages surveyed, Aurepalle in Andra Pradesh state, and Kanzara and Shirapur (both in Maharashtra state). In each village, detailed expenditure and income data were collected for 40 randomly sampled households on an annual basis.

For our analysis we need information on both consumption and income aggregates across households and over time. We follow Laczó (2014) and use a consumption aggregate that includes monthly expenditure on food, clothing, services, utilities and intoxicants, such as paan, alcohol and tobacco.[9] The income aggregate contains net income from farming and livestock, labour and transfers from outside the village. All variables are in real and per-adult equivalent units where the same age-gender weights are used as in Townsend (1994). For comparability with other authors, we restrict our analysis to the years 1976-1981 and construct a fully balanced panel.[10]

The ICRISAT villages are poor with the average dweller living well below the \$1 dollar a day poverty line (Table A.2 in Section A.2 of the Online Appendix). On average, daily nondurable consumption per adult equivalent is 0.83 (Aurepalle), 1.10 (Kanzara) and 1.18 (Shirapur) in 1975 rupees, which is equivalent to 0.48, 0.63 and 0.68 in 2016 US dollars respectively. Income is about twice as high, and the difference between income and consumption might be

8. The ICRISAT panel data set has been used to test the Pareto-efficient risk sharing model with homogenous preferences (Townsend, 1994), with decreasing relative risk aversion (Ogaki and Zhang, 2001) and with heterogenous risk preferences (Mazzocco and Saini, 2012). It has also been used to test the dynamic limited commitment model with homogenous preferences (Ligon et al., 2002) and with heterogeneous risk preferences (Laczó, 2014).

9. We thank Sarolta Laczo for making her version of the data available to us.

10. See Morduch (1991) and Ravallion and Chaudhuri (1997) for a detailed discussion of measurement issues in the full ICRISAT panel and revisions to the data.

Table 1. Parametric and non-parametric estimates of the degree of risk sharing

	Aurepalle	Kanzara	Shirapur
	(1)	(2)	(3)
Panel A: non-parametric			
$\frac{Var_{dc}}{Var_{dy}}$	0.30	0.56	0.33
	(0.052)	(.180)	(.074)
Panel B: OLS, dependent variable Δ *ln of aeq. consumption*			
Δ ln of aeq. income	.206	.222	.169
	(.061)***	(.071)***	(.059)***
Obs.	170	185	155
No. of households	34	37	31

Note: Panel A of the table shows the variance of consumption growth divided by that of income growth in Aurepalle, Kanzara and Shirapur. The measure is conditional on changes in aggregate resources. Panel B shows the results from a regression of consumption growth on income growth in each of the three villages, where consumption and income are demeaned period-by-period before the regression. Standard errors in parentheses are clustered at the household level.

accounted for by durables consumption, investment in livestock and housing, but also measurement error.

There is strong evidence of consumption smoothing. In Panel A of Table 1, we report a first summary measure for the relative smoothness of consumption and income, namely the variance of consumption growth as a proportion of the variance of income growth Var_{dc}/Var_{dy}, after partialling out changes in village resources. We focus on the joint distribution of income and consumption *growth* as it removes unmodelled constant sources of heterogeneity. In all three villages, consumption smoothing is strong, though far from perfect with the variance of consumption relative to income ranging from 0.30 in Aurepalle to 0.56 in Kanzara.

In Panel B of Table 1, we report the coefficient estimates for the following regression

$$d\tilde{c}_t^j = \alpha + \beta_{dcdy} d\tilde{y}_t^j + e_t^j \tag{25}$$

where $d\tilde{c}_t^j$ and $d\tilde{y}_t^j$ denote the growth rate of adult-equivalent consumption and income, both demeaned with respect to the time dimension equivalent to a time fixed effect specification, and e_t^j is an error term.[11] Again, the first-differences specification removes all time-invariant unobserved heterogeneity at the household level. The coefficient β_{dcdy} is a second measure of the degree of insurance, extensively studied in previous work (Townsend, 1994; Laczó, 2014), measuring the share of individual income movements, which passes through to consumption on average. The coefficients on the growth of adult-equivalent income imply that a 1% change in income leads to roughly a 0.2% change in consumption. We take this effect, which is fairly uniform and significant across all three villages, and the smoothness of consumption growth relative to income growth as strong evidence for substantial consumption risk sharing.[12]

In Figure 2, we plot the joint distribution of the residual consumption and income growth (after time demeaning). The data shows little sign of the asymmetry suggested by the limited commitment mechanism in equation (8) and illustrated in the context of village-level risk sharing in the ID model in Figure 1: neither the variance of consumption nor the response of consumption to income

11. We demean the data in this way to make our results comparable to the empirical practice of partialling out aggregate income movements in risk sharing regressions (see e.g. Deaton (1990), Cochrane (1991), Ravallion and Chaudhuri (1997) or Laczó (2014)), and robust to any correlation in individual incomes not captured by the assumption of independent individual incomes. As Laczó (2014) points out, the correlation of incomes across individuals in the three villages is positive, but small. Nevertheless, we decide to be conservative and condition on movements in aggregate village income. Note that conditioning may affect the estimates of β_{dcdy} in the presence of preference heterogeneity when income of less risk-averse households comoves more strongly with aggregate income, as assumed by Mazzocco and Saini (2012). For these reasons, we also estimate the data moments without time-demeaning in Section A.11 in the Online Appendix. The results are very similar.

12. That income exceeds consumption suggests that there may be additional intertemporal mechanisms, and Lim and Townsend (1998) show in their comprehensive study of financial instruments in the ICRISAT data that accumulating crop inventory and currency, but not livestock and other real capital assets, contribute to consumption smoothing. However, the authors also note that the degree of smoothing is too strong to be attributed entirely to saving and borrowing, and conclude that the ICRISAT villages "... appear to be economies in which there is nontrivial social interaction along insurance lines."

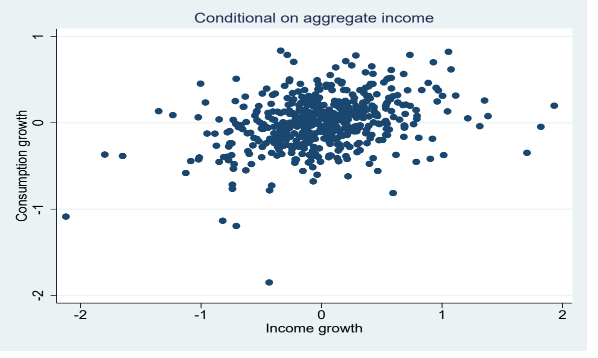

FIGURE 2. Consumption growth and income growth in the ICRISAT dataset

Note: The figure shows a scatter plot of consumption and income growth for households in Aurepalle, Kanzara and Shirapur, where both measures are time-demeaned.

look dramatically different as households move from negative to positive income growth.

We also test the impression of symmetry more formally using two summary moments. First, Panel A of Table 2 reports the relative volatility of consumption and income growth for households with increasing vs non-increasing income. Consumption growth of households that experience positive income growth is, if anything, less volatile than that of those who experience negative or zero income growth. Specifically, the point estimate of their difference (scaled by the variance of income growth to lie between 0 and 1), which is the moment we use in the estimation of the theoretical models in Section 4, is always negative, and for Aurepalle and Kanzara we cannot statistically reject the hypothesis of symmetry (corresponding to a zero difference) at usual levels of confidence.

In panel B of Table 2, we report a second measure of the asymmetry, based on the following regression that aims to capture non-linearities in the conditional mean function

$$d\tilde{c}_t^j = \alpha + \beta_{dcdy} d\tilde{y}_t^j + \beta_{dcdy|dy>0} d\tilde{y}_t^j \times \mathbb{I}_{dy>0} + \gamma \mathbb{I}_{dy>0} + e_t^j \tag{26}$$

where $\mathbb{I}_{dy>0}$ is a dummy variable that takes value 1 when income growth is positive, and 0 otherwise, and where we demean both consumption and income

TABLE 2. Parametric and non-parametric estimates of the consumption response to income rises and falls

	Aurepalle	Kanzara	Shirapur
	(1)	(2)	(3)
Panel A: non-parametric			
$\frac{Var_{dc\|dy>0} - Var_{dc\|dy<=0}}{Var_{dy}}$	-0.08	-0.25	-0.16
	(0.068)	(.243)	(.086)
Panel B: OLS, dependent variable Δ ln of aeq. consumption			
Δ ln of aeq. income	.441	.384	.176
	(.080)***	(.137)***	(.092)*
$\mathbb{I}_{\Delta \text{ln of aeq. income}>0}$	-.035	-.090	.026
	(.062)	(.053)*	(.070)
Δ ln of aeq. income $\times$ $\mathbb{I}_{\Delta \text{ln of aeq. income}>0}$	-.413	-.141	-.053
	(.137)***	(.174)	(.095)
Obs.	170	185	155
No. of households	34	37	31

Notes: Panel A reports the difference in the variance of consumption growth for those experiencing income growth and those experiencing income losses, scaled by the variance of income growth. The measure is conditional on changes in aggregate resources. Panel B shows the results from a regression of consumption growth on income growth, a dummy for households with rising income, and its interaction with income growth. Consumption and income are demeaned period-by-period before the regression. Standard errors in parentheses are clustered at the household level.

period-by-period before the regression.[13] The coefficient $\beta_{dcdy|dy>0}$ provides a second measure of the asymmetry of the joint distribution, by capturing the non-linearity in the conditional mean function of consumption growth around zero income growth. Similar to the previous result, the data features an association of consumption and income growth whose point estimate is smaller for households with rising income, and again, the difference is not statistically different from

13. Demeaning is equivalent to including time dummies in a linear regression of consumption growth on income growth. It amounts to a slight difference, however, when we allow for non-linearities in the association of income and consumption growth in Panel B. Specifically, inclusion of a full set of time dummies would identify the non-linearity only from within-period differences in (already-demeaned) income growth greater than zero. Since our theoretical model does not allow the non-linearity to differ across time, we opt to retain the between-period variance for the identification of the non-linearities.

zero in two of the three villages, in this instance Kanzara and Shirapur. The difference in regression coefficients in Aurepalle, in contrast, is more strongly negative and statistically different from zero.

Both sets of results point in the same direction: there is no evidence of positive asymmetry where the amount of insurance obtained for negative income growth is higher than for positive income growth. On the contrary, the asymmetry displayed in the data is negative (but imprecisely estimated).

4. Results

This section structurally estimates the CD model of Section 2.2 and compares its quantitative implications to data from the three ICRISAT villages, as well as to estimated versions of the ID and SI comparison models.

4.1. Model estimation

This subsection describes how we estimate the income process and preference parameters, and how we determine group size (for the ID and CD models), and the exogenous interest rate (for the SI model).

4.1.1. Group size in the ID and CD model. We use the ICRISAT data to inform the group sizes we consider in the CD and ID model. For the ID model, we follow standard practice in the literature (Ligon et al., 2002; Laczó, 2014) to compute the model with village size equal to the number of households sampled by the ICRISAT. We set $N = 34$ in Aurepalle, $N = 37$ in Kanzara, and $N = 31$ in the case of Shirapur (in reality, all three villages in the ICRISAT dataset comprised several hundred households at the time of the survey).[14] For the CD

14. Although standard in the literature, the assumption that village size and sample size coincide is not ultimately satisfactory. Allowing group size to be larger than the sample should, however, imply even more extreme values for the asymmetry in the standard model. It would only affect the results of the coalitional deviations model inasmuch as there are stable groups beyond the sample size, which is the maximum we consider.

model, we follow the same logic and assume that households can only form risk sharing arrangements with the $N-1$ other households in the village, and, that they will join the largest group within the village for which a stable insurance contract exists. This implies a village size of $N' \geq N$, the smallest size larger or equal to N that is a multiple of the maximum stable group size n^{max} in the CD model (see Section 2.3). Since below we estimate n^{max} to be a small single-digit number, the resulting difference in village size between the ID and CD models is small.

4.1.2. Income process. We estimate separate income processes for each village in the ICRISAT data. We assume that log-incomes of all village members follow an AR(1) process with common persistence parameter ρ

$$y_{it} = \rho y_{it-1} + \varepsilon_{it} \tag{27}$$

where ε_{it} are mean-zero shocks that are identically and independently normally distributed across households.

We estimate ρ and the variance of shocks Var_{ε} from the autocovariance and variance of household incomes $\tilde{y}$ in the ICRISAT data, where $\tilde{y}$ is the residual from a regression of time-demeaned log income on household fixed effects:

$$\rho = \frac{Cov_y}{Var_y}$$

$$Var_{\varepsilon} = Var_y(1-\rho^2) \tag{28}$$

Note that the benchmark version of our model, where households only differ in their income and consumption realisations ex post, abstracts from any ex-ante heterogeneity. This is why we partial out household-specific fixed effects in the estimation of the income processes, to make our results robust to errors in adjusting for household size, and to permanent differences in household incomes (which would otherwise be interpreted as persistent shocks around homogeneous mean income). Also, we time de-mean the data in order to partial out aggregate movements in village income. The Online Appendix, however, reports the results when estimating the income process both without demeaning (Section A.11) and without household fixed effects (Section A.14), and when we allow for ex-ante heterogeneity of households in incomes (in Section A.8) and preferences (in Section A.13).

TABLE 3. Estimated income processes

	Aurepalle	**Kanzara**	**Shirapur**
ρ	0.28	-0.00	-0.18
	(.090)***	(.107)	(.083)**
Var_{α_i}	0.23	0.30	0.36
	(.070)***	(.106)***	(.099)**
Var_{ε}	0.15	0.064	0.11
	(.030)***	(.011)***	(.015)**

Note: The table presents the point estimates for the persistence parameter ρ and the shock variance Var_{ε} for the AR(1) process (27), as well as the variance of household fixed effects Var_{α_i}. Bootstrapped standard errors in parentheses.

Table 3 presents the estimates for the AR(1) parameter ρ and the shock variance Var_{ε} for the three villages. Incomes have low positive persistence in Aurepalle ($\rho = 0.28$), are serially uncorrelated in Kanzara, and have small negative serial correlation in Shirapur ($\rho = -0.18$). Aurepalle has the most volatile income shocks of the three villages. Table 3 also reports the variance of the fixed effects, Var_{α_i}, which accounts for a sizeable fraction of the total variance of individual incomes.

In order to use the estimates of ρ and Var_{ε} for the quantitative solution of our model, we approximate y_{it} as a Markov process with three support points using the Rouwenhorst (1995) method. The Online Appendix A.1 reports the transition matrices.

4.1.3. Preferences. We assume that per-period utility is of the constant-relative-risk-aversion type

$$u(c_t) = \frac{c_t^{1-\sigma} - 1}{1 - \sigma} \tag{29}$$

As common in limited-commitment economies with strong insurance (see e.g. Laczó (2014), and Section A.4 in the Online Appendix), we cannot separately identify the discount factor δ and the coefficient of relative risk aversion σ from the joint allocation of consumption and income. We therefore normalise σ to 1 (log-preferences), and estimate the discount factor by matching moments in a long simulation of our model to those in ICRISAT data. Section A.3 in the Online Appendix explains the estimation procedure in detail.

Informed by Figure 1, we choose four target moments of the joint distribution of individual consumption and income growth that have a close link to intuitive features of the models such as the degree of risk sharing and the degree of asymmetry in the implied reaction of consumption to positive and negative income shocks. Apart from the commonly used average slope of the conditional mean (the regression coefficient β_{dcdy} in (25), as in for example Townsend 1994), we also choose as a target the relative variance of consumption and income growth $Var_{dc}Var_{dy}$. Whenever the model is non-linear like ours, this is an additional important summary measure for the degree of insurance. The estimates of these moments in the ICRISAT data have been reported in Table 1. To make the simulated and data moments more comparable, recall that we have purged the data moments of aggregate time variation and individual time-invariant heterogeneity.

To capture the asymmetry suggested by equation (8), Figure 1 suggests to focus on the non-linearity in the conditional mean, and in the conditional variance, around 0 income growth. We therefore consider first the difference between the variance of consumption growth of households that experience positive income growth and that of those that do not, $(Var_{dc|dy>0} - Var_{dc|dy\leq 0})/Var_{dy}$ (scaled by the variance of income growth); and second the difference in the regression coefficient of consumption growth on income growth for households with rising and non-rising income, $\beta_{dcdy|dy>0}$ and $\beta_{dcdy|dy\leq 0}$.[15] The estimates of these moments in the ICRISAT data have been reported in Table 2.

4.2. Model estimates

Table 4 presents the moments of interest when the discount factor δ is estimated to target the two moments that summarize the average degree of insurance in the three villages, Var_{dc}/Var_{dy} and β_{dcdy}. To interpret the estimates, it is useful to recall the role of the discount factor in the limited commitment environment:

15. For the calculation of these moments, we group periods of constant income together with those of falling income. Relative to an alternative procedure that leaves periods of constant income aside in the moment-calculation, this does not change the substance of the results.

Table 4. Preferences estimated to target degree of risk sharing - 2 moments

	Aurepalle				**Kanzara**				**Shirapur**			
	Data	**CD**	**ID**	**SI**	**Data**	**CD**	**ID**	**SI**	**Data**	**CD**	**ID**	**SI**
n		4.00	34.00			4.00	37.00			4.00	31.00	
δ		0.97	0.83	0.92		0.98	0.87	0.93		0.96	0.82	0.92
s.e.		0.01	0.01	0.01		0.01	0.02	0.01		0.01	0.02	0.02
σ		1.00	1.00	1.00		1.00	1.00	1.00		1.00	1.00	1.00
$\frac{Var_{dc}}{Var_{dy}}$	0.30	0.23	0.22	0.18	0.56	0.24	0.18	0.14	0.33	0.23	0.13	0.12
β_{dcdy}	0.21	0.23	0.33	0.34	0.22	0.23	0.30	0.28	0.17	0.22	0.25	0.24
$\frac{Var_{dc\mid dy>0}-Var_{dc\mid dy\leq 0}}{Var_{dy}}$	-0.08	0.00	0.29	0.10	-0.25	0.00	0.23	0.07	-0.16	0.00	0.17	0.06
$\beta_{dcdy>0} - \beta_{dcdy\leq 0}$	-0.41	-0.00	0.48	0.24	-0.14	0.00	0.43	0.23	-0.05	-0.00	0.36	0.24
Model criterion		1.95	6.56	10.21		3.31	5.72	6.07		2.58	8.28	8.80

Note: For each of the three ICRISAT villages, the table shows four key moments summarising the joint distribution of consumption and income growth in the survey ("Data", in the first column for each village), and in simulations of the CD model (in the second column for each village), the ID model (third column) and the SI model (fourth column). For the simulated model solutions, the table also presents the size of the insurance groups n (in the case of the ID and CD models) and the discount factor δ, which is chosen to minimise the sum of differences between the first two moments (Var_{dc}/Var_{dy} and β_{dcdy}) predicted by the models and those observed in the data, weighted by the inverse variance of the data moments calculated using a bootstrap procedure. The estimated preference parameters are those that minimise this criterion on a grid of $\delta \in [0.5, 0.99]$. The last line reports the value of the criterion at the estimated parameters.

Table 5. Preferences estimated to target degrees of risk sharing and asymmetry

	Aurepalle				Kanzara				Shirapur			
	Data	**CD**	**ID**	**SI**	**Data**	**CD**	**ID**	**SI**	**Data**	**CD**	**ID**	**SI**
n		4.00	34.00			4.00	37.00			4.00	31.00	
δ		0.97	0.89	0.93		0.98	0.87	0.94		0.96	0.88	0.94
s.e.		0.01	0.01	0.01		0.00	0.01	0.01		0.01	0.01	0.01
σ		1.00	1.00	1.00		1.00	1.00	1.00		1.00	1.00	1.00
$\frac{Var_{dc}}{Var_{dy}}$	0.30	0.23	0.04	0.12	0.56	0.24	0.18	0.10	0.33	0.23	0.02	0.07
β_{dcdy}	0.21	0.23	0.11	0.26	0.22	0.23	0.30	0.23	0.17	0.22	0.07	0.16
$\frac{Var_{dc\|dy>0}-Var_{dc\|dy\leq 0}}{Var_{dy}}$	-0.08	0.00	0.07	0.07	-0.25	0.00	0.23	0.05	-0.16	0.00	0.03	0.04
$\beta_{dcdy>0} - \beta_{dcdy\leq 0}$	-0.41	-0.00	0.20	0.21	-0.14	0.00	0.43	0.20	-0.05	-0.00	0.13	0.17
Goodness of fit		12.28	52.76	38.66		5.07	20.41	11.93		6.14	27.17	22.24

Note: For each of the three ICRISAT villages, the table shows four key moments summarising the joint distribution of consumption and income growth in the survey ("Data", in the first column for each village), and in simulations of the CD model (in the second column for each village), the ID model (third column) and the SI model (fourth column). For the simulated model solutions, the table also presents the size of the insurance groups n (in the case of the ID and CD models) and the discount factor δ, which is chosen to minimise the sum of differences between all four moments predicted by the models and those observed in the data, weighted by the inverse variance of the data moments calculated using a bootstrap procedure. The estimated preference parameters are those that minimise this criterion on a grid of $\delta \in [0.5, 0.99]$. The last line reports the value of the criterion at the estimated parameters.

since deviation delivers higher mean consumption in earlier periods at the price of eternally higher consumption volatility, higher discount factors, like higher risk aversion, deter deviation and increase risk sharing. Importantly, the estimated value of the discount factor is conditional on the normalisation of relative risk aversion to 1 (log-preferences). A normalisation to higher risk aversion would thus deliver lower estimated discount factors. In other words, when interpreting the estimation results below, the focus should be on the relative values across models and estimation criteria, not the absolute level of the discount factor.[16]

To match the high degree of insurance in the ICRISAT villages, the CD model estimates discount factors between 0.96 and 0.98 in the three villages. It requires such large discount factors both because the outside option in the CD model is very attractive and because its maximum sustainable insurance groups are substantially smaller than the ICRISAT sample, comprising 4 households in all villages. At these group sizes, close-to-full insurance is necessary to match the data. The estimated discount factors that match the two target moments in the ID model are much lower, ranging from 0.82 to 0.87. This is explained by the fact that the ID model trades off its larger (exogenously given) group size against a lower discount factor. Estimated discount factors in the SI model lie between the ID and CD model, ranging from .92 to .93. The standard errors of these estimates are generally small, but somewhat larger in the CD model (where insurance is close to perfect in small groups, and small changes in the discount factor thus have smaller effects on the predicted moments).

All three models produce a reasonably good fit of the degree of insurance. The CD model predicts an average association of consumption and income growth, as measured by the regression coefficient β_{dcdy}, slightly stronger than in the data, and somewhat underpredicts the relative volatility of consumption growth (most strongly in the village of Kanzara). The ID and SI models also produce estimates that are reasonably close to the target moments, but their fit is worse than in the CD model.

16. Figure A.1 in the Online Appendix confirms that the two preference parameters are not separately identified. It shows that, within the range of discount factors that are consistent with the observed degree of insurance, for every value of the discount factor δ there is a value of risk aversion σ that produces the same minimized model criterion, and the same corresponding moments, for both models.

Importantly, the CD model matches the average degree of insurance while predicting an approximately symmetric consumption-income growth distribution: the coefficients summarizing the asymmetry are both close to 0. In contrast, both comparison models predict strong counterfactual asymmetries in the conditional mean function (as measured by the difference in regression coefficients $\beta_{dcdy>0} - \beta_{dcdy\leq 0}$), and in the conditional variance function (as indicated by the difference in variances $(Var_{dc|dy>0} - Var_{dc|dy\leq 0})/Var_{dy}$). The asymmetry in the SI model is, perhaps, more surprising, since the simplest version of self-insurance, the permanent income hypothesis (PIH), would predict the change in consumption to equal the change in permanent income, and thus a symmetric reaction to income rises and falls. As explained in Krueger and Perri (2005), however, with borrowing constraints consumption responds more to income changes when asset buffers are smaller. Since negative income shocks reduce assets (as households dis-save) and make positive income growth more likely (as incomes are predicted to revert to their means), positive income shocks occur more often at low asset values, and are thus associated with larger consumption increases.

What do the estimated versions of the three models imply for the joint distributions of consumption and income? We examine this in Figure 3 with the help of scatter plots for the village of Aurepalle. The middle panel, depicting the ID model, is similar to Figure 1. Particularly, it features a similarly pronounced kink in both the conditional means and variances of consumption growth around zero income growth. The distribution generated by self-insurance, in the right panel, has an asymmetry that is somewhat smaller than that in the ID model with a less pronounced heteroscedasticity. Finally, the left panel, depicting the CD model, shows a homoscedastic distribution around a linear conditional mean function. This is in line with the absence of asymmetry in its estimates in Table 4 and the joint distribution of income and consumption growth in the ICRISAT villages depicted in Figure 2 in Section 3.

Table 5 shows that, when we include the two asymmetry moments in our estimation criterion, the estimates of the CD model are effectively unchanged. At a higher estimated discount factor δ the ID model now substantially overpredicts the degree of insurance while retaining a counterfactual asymmetry. The estimates of the SI model are also of increased discount factors and insurance, but less so than in the ID model, implying a smaller reduction in the asymmetry.

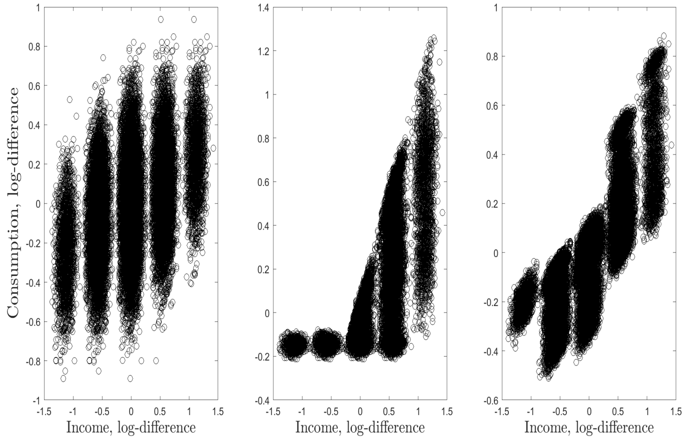

FIGURE 3. Consumption and income growth in general equilibrium

The figure shows scatter plots of consumption and income growth from a simulation of the CD, ID, and SI economies, for the income process and preferences estimated for Aurepalle in Tables 3 and 4, respectively. The figure plots residuals from a regression that controls for movements in aggregate resources.

In line with these results, the goodness of fit of the two comparison models, already worse in Table 4, further deteriorates both in absolute terms and relative to the CD model when evaluated using all four moments of interest.[17]

4.3. Inspecting the mechanism: the role of group size

The two limited-commitment models rely on very different mechanisms to arrive at their prediction of the observed degree of risk sharing. The ID model

17. It is tempting to compare the corresponding measures in the final rows of Tables 4 and 5 to a χ^2 distribution. This would indicate substantially higher p-values in the CD model than for the comparison models, but is valid only under the strong assumption of independent moment conditions. Ultimately, the aim of our exercise is, however, not to reject, or not, any of the, still very stylized, models, but to highlight their different implications for the structure of consumption risk sharing and their ability to capture key moments of the data.

predicts moderate degrees of insurance taking place at the village level. The CD model predicts strong insurance among a smaller number of households, which translates into moderate degrees of insurance at the village level.[18] Importantly, this strong insurance at the group level reduces the increase in asymmetry when group size increases. This role of group size and the degree of insurance for the asymmetry is illustrated in Figure 4 which varies group size exogenously in the ID model. The figure depicts, as a measure of insurance, $1 - \beta_{dcdy}$ (the dashed line) and, as a measure of asymmetry in consumption responses, the difference between the regression coefficients $\beta_{dcdy|dy>0}$ and $\beta_{dcdy|dy\leq 0}$ (the solid line) as a function of group size (along the bottom axis) and for two values of δ implying moderate (top panel) and strong (bottom panel) insurance.

As expected, the degree of insurance (as indicated by the dashed lines) is lower when agents are more impatient (in the top panel). For a discount factor of 0.86 (the top panel), the model matches the observed insurance in Aurepalle when the insurance group comprises the entire village (of 34 households), and insurance within the village is only partial. When households are more patient ($\delta = 0.91$, in the bottom panel), in contrast, insurance is stronger and matches that in the data when insurance is essentially perfect within smaller groups that consist of only 5 households. Asymmetry is 0 when insurance groups are pairs, and increases monotonously with village size, but is substantially smaller when insurance is strong.[19]

18. This feature of strong insurance in small insurance groups within larger villages makes the equilibrium similar to the model environment in Heathcote et al. (2014), who assume unconstrained complete markets within islands that only trade bonds among each other. With random walk island-specific shocks, and many households within islands, this results in no (perfect) risk sharing across (within) islands. The similarity is not perfect, however, since the endogenous groups we estimate features a small number of households, and we assume that coalitions cannot trade any assets among each other.

19. There is, potentially, an additional, more mechanical reason for lower asymmetry in the alternative model where, as explained at the beginning of this section, we keep the number of village members approximately equal to that in the standard model by simulating multiple insurance groups. The resulting average village income and consumption, which we partial out, is less than perfectly correlated with average incomes in the insurance group. Thus, with multiple risk sharing groups, conditioning

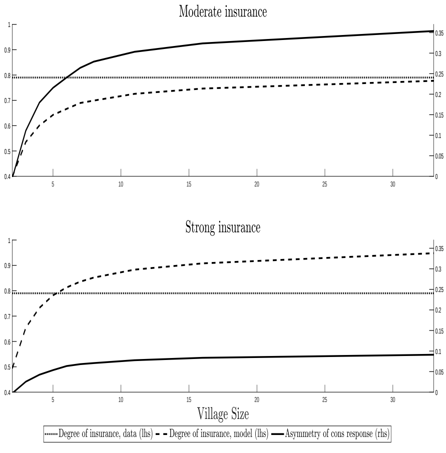

FIGURE 4. Insurance and asymmetry in the standard model

The figure presents results from a simulation of the ID model with rising group size. The figure presents two key moments that summarize the asymmetry and the degree of insurance as a function of village size (along the bottom axis) and for two values of δ implying moderate (top panel) and strong (bottom panel) insurance: first, one minus the regression coefficient of consumption growth on income growth $(1 - \beta_{dcdy})$ as a measure of the degree of insurance (the dashed line, indicated on the left axis); and second, the difference between the regression coefficients of consumption growth on income growth for households with rising and non-rising income, $\beta_{dcdy|dy>0}$ and $\beta_{dcdy|dy\leq 0}$ (the solid line, indicated on the right axis).

Figure 4 therefore shows that the CD model achieves its superior fit largely because of its prediction of strong insurance in smaller group sizes rather than a superior performance than the ID model at given group size. This naturally gives rise to the question if the CD model of endogenous sustainable group size is preferable to a version of the ID model with exogenously small groups. We examine this question in Section 5.

---

on village variables leaves a group-component in household-level variables that makes the observed data more symmetric. Since the treatment of the simulated data follows directly from the standard conditioning we apply to the empirical data, this differential effect does not imply, in our view, any inconsistency. As a robustness exercise, however, we repeat the simulated method of moments estimation on the unconditional moments in Section A.11 of the Online Appendix.

4.4. Further analysis and sensitivity

A separate Online Appendix contains further analysis. In Section A.5, we demonstrate the accuracy of the approximation underlying our benchmark estimates relative to exact solutions of the CD model. In Section A.5.1, we first compare the results from that approximation to the exact solution of the model for the small groups where the latter is feasible. For groups of three and four, the moments predicted by the approximation are very close to the exact ones over the relevant range for the discount factor. Thus, conditional on a small group size, we are confident that our approximation is accurate. This confidence rests on the maintained assumption, however, that the approximation indeed accurately identifies sustainable groups. While we cannot solve the exact model in its general version for larger group sizes, we can identify sustainable groups in the exact model whenever perfect insurance is sustainable. This is relevant in particular as insurance is indeed close to perfect in our benchmark estimation. In a second robustness check in Section A.5.2, we therefore report the maximum sustainable groups under a maintained norm of full insurance. The sizes we estimate in this alternative way are identical.

Although simpler, both the approximation and the exact CD model imply equilibrium transfer rules that may be too complex to implement in reality. Thus, we examine in Section A.6 whether there is a simpler rule-of-thumb implementation of the constrained-efficient contract that keeps the main features of the full model but is easier to put into practice. Within the set of such simpler rules we estimate there to be full risk sharing in small groups, in contrast to our benchmark estimates of strong but not perfect risk sharing.

We then turn to a (partial) examination of how sensitive the finding of small equilibrium groups in the CD model is to the core assumptions in the benchmark estimation. First, we examine alternative specifications for the outside option in the CD model and second, we allow for more heterogeneity among households along a number of dimensions. In Section A.7, we show that the results are unaffected with a different choice of outside option for the rest of the group in the CD model, namely constrained-optimal rather than first-best risk sharing. Second, when deviations are made less attractive by forcing members of deviating subgroups to consume their income in the period of deviation (as, for example, in Genicot and Ray (2003)), the degree of insurance improves. This is counteracted by a fall in the estimated discount factor δ that leaves the predicted moments largely unaffected. In Section A.8, we allow for households to have ex-ante heterogeneous income processes. In Section A.9, we depart from the assumption that all members of a deviating coalition are (initially) treated equally after deviation. Instead unequal treatment is possible in the sense that the deviating coalition inherits the relative Pareto weights of the existing larger group. In both cases, we estimate a high degree of insurance in small insurance groups and a good match between model and data, just as in the benchmark estimation.

Next, we report on four robustness checks that show that the finding of a strong positive asymmetry in the standard ID limited commitment model is not dependent on particular choices for the numerical solution, the model environment, or moments to target. First, section A.10 illustrates that the asymmetry does not depend on how the income process is discretized and that it is unchanged (both in the ID and the CD model) when the income process has more than the three support points we consider in the benchmark solution. Second, section A.11 shows that the results are unaffected when we do not partial out aggregate variation in the targeted data and simulated moments. Third, Section A.12 reports results when we freely estimate the persistence parameter ρ in the CD and ID models to target the joint distribution of consumption and income in the data. Even with a serial correlation that is counterfactually negative relative to the data, the standard model is not able to simultaneously predict a symmetric distribution of consumption and income growth and a realistic observed degree of insurance. Finally, one might worry that the asymmetry is a consequence of preference homogeneity. In contrast, with preference heterogeneity, the right mix of approximately risk-neutral and highly risk-averse households, experiencing perfect and zero income-consumption comovement respectively, can always deliver the right average comovement and symmetry in our moments-based approach. We show, however, that the asymmetry is little changed when we allow a simple form of heterogeneity in risk-preferences in the ID model (Section A.13).

A final set of checks examines robustness of the results with respect to the treatment of unmodelled variation in the data when fitting the model to the data. First, section A.14 shows that our conclusions are also unchanged when estimating the income process from the ICRISAT data without conditioning on household fixed effects to control for unmodelled heterogeneity across households. This is important because, in short samples like ours, the household-level fixed effects that we include in the estimation of the income process introduce a downward bias in the AR(1) coefficient ρ. Without fixed effects, in contrast, both the estimated persistence parameter ρ and the variance of income shocks Var_{ε} are necessarily larger. This implies offsetting effects on the outside option of autarky: higher persistence makes autarky more attractive for high-income individuals, while higher volatility makes it less attractive. Together with an increase in the estimated discount factor, this keeps the prediction from the CD model, of strong insurance in small groups, unaffected, and the moments estimated for both the CD and ID models close to those in the benchmark case (apart from a reduction of the predicted group size in Kanzara to 3 households).

Second, section A.15 estimates the three models with measurement error in incomes and consumption. Measurement error in consumption is estimated to be substantial in all three models, and brings the predicted volatility of consumption growth Var_{dc}/Var_{dy}, which was counterfactually low in all benchmark estimates, in line with the data (thus essentially removing it from the estimation criteria). Measurement error in incomes attenuates the regression coefficient $\beta_{dcdy|dy}$ (to counterfactually low levels particularly in the ID model) and 'blurs' the distinction between income increases and declines in measured data. This strongly improves the fit of the SI model. In the CD and ID models, in contrast, some estimates with income measurement error are not well identified and should therefore be treated with caution.

5. Risk sharing in small insurance groups

In this section, we provide evidence beyond the joint distribution of individual consumption and income that supports risk sharing in small insurance groups in our dataset. First, we review the empirical literature on mutual insurance networks and groups to show that there is ample evidence across a variety of contexts that insurance takes place in smaller subgroups within communities. Second, we show that the pattern of pairwise consumption correlations of households in the ICRISAT villages is inconsistent with risk sharing taking place at the village level, but strongly supports risk sharing in small groups. Third, we show that the CD model can generate negative asymmetries in the response of consumption to income and discuss what data would be required in order to use this property to distinguish between endogenously and exogenously small groups, something which we cannot do in the context of the ICRISAT data.

5.1. Existing evidence for small risk-sharing groups

Our prediction of strong insurance among smaller groups of households is in line with a large literature in development economics that shows how risk sharing often takes place in smaller groups. Most relevant to our argument is the literature that maps relevant insurance networks by asking households to identify insurance partners they rely on in times of need (Fafchamps and Lund, 2003; De Weerdt, 2004). Fafchamps and Lund (2003) find that households in the Philippines make and receive transfers from an average of 5 other households. Our own calculations using data on social networks in South India (see Banerjee et al. (2013)) show that a household is connected to an average of just over 3 other households for the purposes of mutual help in times of need.

5.2. Pairwise consumption correlations in the ICRISAT data support small risk-sharing groups

The original ICRISAT data does not allow for a mapping of insurance groups, since it does not include information on how many villagers a household cooperated with in situations of need, and/or who they received transfers from, or gave transfers to, in a given year. To provide additional evidence of small risk-sharing groups in this data set, we therefore follow an indirect approach to identify groups that builds on the intuition that consumption comovement should be stronger within than across insurance groups (see for example Ligon (2004)). This suggests that the distribution of bilateral consumption correlations within a village should be informative about the size of insurance groups.

To illustrate the relationship between bilateral consumption correlations and insurance group size, we first simulate the distribution of bilateral consumption correlations in a village that is partitioned into equal-sized insurance groups of varying sizes. Specifically, using the income process in the village of Kanzara as input, we solve the ID model for three group sizes, namely $n = 4$, which corresponds to our estimated group size in the CD model, $n = 37$, the sample size in Kanzara and insurance group size used in the ID model, and a village-level

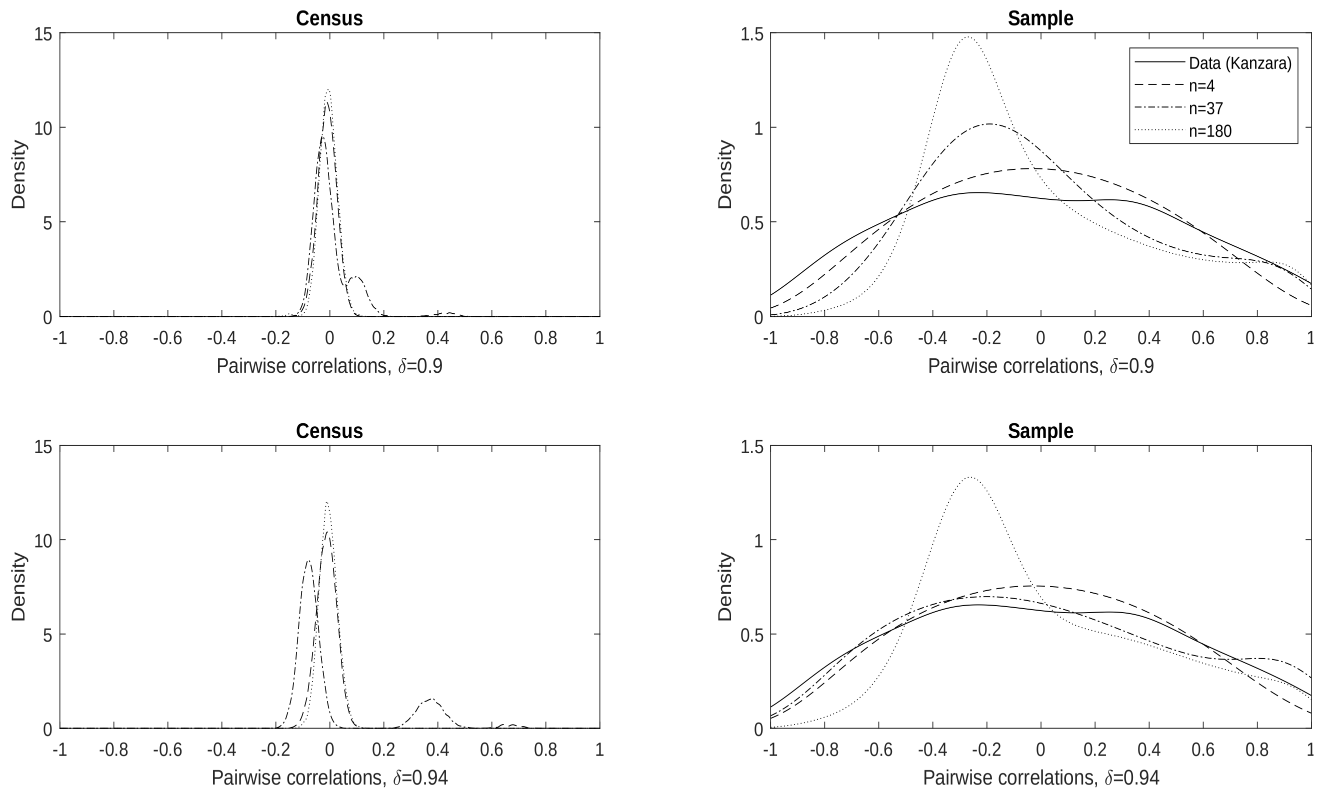

FIGURE 5. Model-implied bilateral consumption correlations within a village

The figure shows kernel estimates of the density of pairwise consumption correlations in a village whose size and income process correspond to Kanzara, and where households are divided into insurance groups of size 4 (dashed line), 37 (dashed-dotted line) and 180 (dotted line). For the panels in the left-hand column of the figure ("Census") the data are generated by the ID model from a large number of simulated panels with $n = 180$ and a large number of time periods, controlling for time fixed effects. For the right-hand column ("Sample") the data are generated by a large number of 6-year samples of size $n = 37$, controlling for time fixed effects. The figure also plots a kernel estimate of the density of pairwise consumption correlations (after partialling out time fixed-effects) in the ICRISAT data for Kanzara (solid line).

group equal to $n = 180$ households, the village size in Kanzara.[20,21] We solve the model for discount factors $\delta = 0.9$ and $\delta = 0.94$, in order to generate moderate and high levels of insurance. The model solutions for a given n and discount factor δ are used to simulate a long sequence ($T = 1,000$) of income and consumption for the N' households in the village, who are divided into k groups of n members (where k is again the smallest integer such that $k \times n = N' \geq 180$, the village size in Kanzara). From these consumption sequences, we then compute (after time-demeaning) the $N' \times N'$ matrix of 'census' bilateral consumption correlations.

The left column of Figure 5 plots kernel estimates of the density of the 'census' bilateral consumption correlations when the village is divided into groups of size 4 (dashed line), of size 37 (dashed-dotted line), or the insurance group comprises the whole village (dotted line). For groups smaller than the village, the distribution of the bilateral correlations is bimodal, as correlations between villagers in different groups cluster around zero (or slightly to the left of it for groups of size 37), but are positive for those within the same group.[22] The correlation for those sharing membership of a risk-sharing group is increasing in the discount factor (as a higher discount factor implies more insurance), and decreasing in group size (as group-income is more volatile in smaller groups whose consumption thus comoves more strongly relative to idiosyncratic consumption and income movements). When the insurance group coincides with the village, in

20. We focus on Kanzara in this analysis, because it is the village with the highest ratio of sampled households (37) to households living in the village at the time of the survey (180), making it easier to identify risk-sharing groups within the village (see Binswanger and Jodha (1978)).

21. Using the whole income distribution as a state variable becomes infeasible for the village-level model with $n = 180$. We thus use a procedure similar to that in Laczó (2014), based on a discretized support of aggregate village income, rather than the full support of all possible aggregate income realizations. This, however, makes the degree of insurance at given preference parameters difficult to compare to those at smaller group sizes.

22. The mode for $n = 37$ is slightly to the left of zero because the correlation of the time-demeaned consumption changes (always) depends negatively on the correlation between own and village consumption, and this correlation becomes sizeable when the discount factor is high and group size large relative to village size.

contrast, time fixed effects capture the entire variation in group-level resources, leaving only idiosyncratic consumption movements that imply a zero correlation in the village.

The simulations show that the patterns of bilateral consumption correlations differ strongly depending on whether risk sharing takes place at or below the village level. To compare the model-generated pairwise correlations to the data for Kanzara, however, we need to take account of the fact that the ICRISAT sample is not a village census. We thus need to consider the possibility that households share risk with others in the village that are not in the data sample. Whenever both the sample and risk-sharing groups are small relative to the village population this increases the number of observed 'zero' correlations, corresponding to unconnected households, relative to the case where the sample includes the whole village. Moreover, because we want to compare the entire distribution of bilateral consumption correlations in the model to that in the data (rather than average moments as we did in the benchmark estimation), we evidently need to modify our simulation approach to take account of the short time-dimension in the ICRISAT panel.

To compare model and data, we therefore generate 'sample' bilateral consumption correlations as follows: From the $N' \times 1000$ matrix of consumption sequences simulated from a model with insurance groups of size n, we draw a 6-period panel of 37 households, equivalent to panel length and sample size in Kanzara. From this short sample, we calculate the 37×37 matrix of bilateral consumption correlations and then repeat the sampling 10,000 times to produce two summary statistics: (i) an average matrix of bilateral consumption correlations, which is the mean of the correlation matrices over the 10,000 short samples, and (ii) a minimum Kolmogorov-Smirnov test statistic, which is a measure of the deviation between two densities. The latter is calculated by choosing from the 10,000 short samples that one whose density of the bilateral consumption correlations gives the best fit to the data (as measured by a low Kolmogorov-Smirnov test statistic).[23] This process is repeated 1,000 times and

23. When insurance groups are smaller than the village, the ability of the model to explain the observed correlation pattern crucially depends on picking the same mix of connected and unconnected households as in the ICRISAT sample. Since the latter is unobserved, we focus for each group size on the simulated panel that gives the best fit

the average test statistic and p-value over the 1,000 repetitions, as well as the minimum over all repetitions, are recorded.

The density of the average 'sample' bilateral consumption correlations is plotted in the right column of Figure 5. Here, we also add the density of the bilateral correlations observed in the data, which are calculated after partialling out household fixed effects and time-averages. Compared to the 'census' densities, the model-generated densities from short samples are now much flatter. And the small-group density ($n = 4$) in particular traces out the shape of the density estimated on the data for Kanzara (the solid line) quite closely. Importantly, both the data and simulations for insurance groups of $n = 4$ and $n = 37$ contain many pairs of households whose consumption is strongly negatively and strongly positively correlated. When the insurance group coincides with the village, in contrast, the simulated density does not capture this dispersion, as its mass remains too concentrated around the mean.

Similarly, the deviation between the data and model generated bilateral correlations is smallest for the small group model with $n = 4$, where the test of equality of the distributions has a p-value of 0.99, when we consider the minimum over all permutations (Table 6). In contrast, for a group of $n = 34$, the minimum p-value is 0.79. Looking at the average model criterion and p-value (where we first find the minimum KS test statistic over 10,000 permutations and then take the average of this minimum over 1,000 repetitions) shows that the $n = 4$ model fits the data much more frequently than the $n = 37$ model (an average p-value of 0.91 as compared to an average p-value of 0.155). Risk sharing at the village level is soundly rejected.

The reason that the bilateral consumption correlations generated by small insurance groups fit the data better is two-fold: (i) because the correlation between any two households is calculated on the basis of only six time periods, there is a much wider range of possible correlations, both for households who share membership of a group (varying around a positive average correlation), but especially for those who do not (varying around a zero average correlation). (ii) When risk-sharing groups are small relative to the sample (and village) size, the sample is mostly made up of households whose consumption is not connected

between model and data, assuming that this is the one that comes closest in terms of getting the mix of members and non-members right.

Table 6. Testing for equality of the data and model generated distribution of pairwise consumption correlations

Kanzara			
	CD	**ID**	**ID**
n	4.00	37.00	180
δ	0.94	0.90	0.90
Min. Model criterion	0.0146	0.0248	0.0628
Min. p-value	0.999	0.787	0.009
Avg. Model criterion	0.0209	0.0505	0.0964
Avg. p-value	0.911	0.155	0.000

Note: For the village of Kanzara, the table reports output from a Kolmogorov-Smirnov test that tests for equality of the distribution of bivariate consumption correlations in the data and the distribution generated by a model of risk sharing in groups of $n = 4$ and $\delta = 0.94$ (the estimated group size and discount factor in the CD model), $n = 37$ and $\delta = 0.9$ (the sample size in Kanzara and estimated discount factor in the ID model), and $n = 180$ and $\delta = 0.9$ (the number of households living in Kanzara). The statistics are based on 10,000 samples consisting of 37 households in 6 time periods drawn from the simulated model data. Over these samples, the minimum KS test statistic and the associated p-value are recorded. The table reports the average of these statistics over 1,000 repetitions of this exercise.

through membership in a group. This makes these model-generated densities more similar to the relatively symmetric and dispersed density for Kanzara.

5.3. Comparing the CD model to an ID model with exogenously small groups

If small insurance groups exist in the ICRISAT villages, this begs the question whether they are endogenously small. This is clearly an important question, particularly because the policy implications of the two models may be quite different: in the CD model group size changes endogenously in response to policy changes, while in the ID model, group size does not respond.

Although we think exogenous barriers that limit insurance groups to 4 or 5 households are difficult to motivate, we now compare the CD model to a version of the ID model whose insurance groups are exogenously set to the same small size. Any difference between the two models thus results from their different outside options for constrained subgroups. Instances of constrained subgroups are rare, however, in a setting such as the ICRISAT, where the CD model generates moderate insurance at the village via near-perfect insurance in small groups and enforcement constraints thus seldom bind.

To highlight the differences between the two models, we thus no longer attempt to match the degree of insurance in the ICRISAT data, but instead perform a comparison at lower discount factors, which imply a lower degree of insurance. Specifically, we solve the version of the CD and ID model in Section 2.2 in groups of size 3 on an income grid with three support points. We use the estimated income process for Aurepalle, but set the time preference parameter δ to 0.86 in order to generate frequently binding enforcement constraints in the two models. We use the exact solutions of the models for this exercise, (i) because when we solve the model only for small groups, we do not have to resort to the approximation, and (ii) since the approximation treats the rest of the village in the same way in either model, this blurs some of the differences between the two models that may help us to distinguish them for the same given group size.

In Table 7, we present the results from simulating the target moments related to the average degree of risk sharing, β_{dcdy} and its asymmetry $\beta_{dcdy>0} - \beta_{dcdy\leq 0}$. As expected, the degree of insurance is now much lower in both models. More interestingly, at this lower level of insurance, the CD model, in line with the data, is able to generate a negative asymmetry in the consumption response of income winners and income losers. That is, consumption changes of income winners are *less* responsive to income changes than those of income losers.

While this qualitative difference between the two models is quantitatively small when we examine the joint distribution of income and consumption at the village level, it becomes much larger when we home in on what happens inside an insurance group and restrict attention to those states where the predictions in the CD model differ from those in the ID model, that is when in each insurance group two constrained agents face an unconstrained one. Specifically, in columns 3 and 4 of Table 7, we simulate the moments only for those periods in which all but one member of the insurance group experienced an income increase. This increases the negative asymmetry in the CD model to -0.15, and the difference between the asymmetry moment in the CD and ID model rises to 0.16, more than a third of the degree of risk sharing in the ID model.

To understand how the negative asymmetry in the regression coefficient arises in the CD model, it is useful to recall how consumption in the ID model is determined when $n-1$ agents have a binding constraint. Ligon et al. (2002) show that consumption of a constrained agent depends predominantly on her own income. Hence when $n-1$ agents are constrained, then, from the budget constraint, an income change for the sole unconstrained agent translates also into a strong consumption response.

The same argument holds for an unconstrained agent also in the CD model, but the consumption response of the constrained agents is muted relative to the ID case. Because the outside option now involves risk sharing in constrained sub-groups, consumption of a constrained agent depends mainly on joint income in the constrained sub-group and less on her own income realization. As a result, an income increase of a constrained agent is spread over the members of the constrained sub-group, thus dampening the response of consumption to own income changes.

To be clear, the different outside options in the two models translate first and foremost into a prediction that consumption of constrained agents is less responsive to income changes in the CD model than in the ID model. This prediction is particularly useful if, as is the case in the numerical example shown here, the consumption response to income changes for the constrained in the CD model is so weak that it is smaller than the consumption response of unconstrained agents. In that case, it is possible to distinguish between the two models on the basis of opposite signs (positive in the ID model and negative in the CD model) for the asymmetry in the consumption response to income changes. Of course, this is merely an illustrative example to show that it is possible to distinguish the CD and ID model at given small group size. More work – beyond the scope of this paper – would be needed to establish that these patterns hold more generally.

Given these findings, what kind of data would one need in order to distinguish unequivocally between endogenously and exogenously limited insurance groups? Beyond data on insurance groups and all their members' consumption and income, households would have to be impatient, and insurance groups small. These features would guarantee that (i) there are many instances in which all but one member of the insurance group are constrained, and (ii) a strong relationship between being constrained/unconstrained and income gains and losses. Given these requirements not just on the richness of the data, but also on the parameters of the data generating process, such as preferences and size of insurance groups, the best opportunity for distinguishing the two models at given group size may be given by a lab- or lab-in-the field experiment where many features of the data generating process can be controlled by the researcher.

Table 7. Negative asymmetry in the CD model

	A: Full sample		B: Restricted I	
	ID	**CD**	**ID**	**CD**
β_{dcdy}	0.39	0.47	0.40	0.48
$\beta_{dcdy>0} - \beta_{dcdy\leq 0}$	0.04	-0.06	0.01	-0.15

Note: For the village of Aurepalle, the table reports two moments summarising the joint distribution of consumption and income growth and its asymmetry in simulations of the CD and ID models with groups of size 3 and $\delta = 0.86$. In Panel A, we report the moments calculated from simulated data of $N' = 36$ households in 12 insurance groups over 6,000 time periods. In Panel B, we use only those periods in which all but one member of the risk-sharing group have experienced income increases.

6. Conclusion

In this paper, we have argued to replace the 'village', or in fact any other exogenous risk sharing group in poor agricultural communities, with a concept of endogenous groups of mutual insurance. For this, we have proposed a quantitative model of dynamic risk sharing with limited commitment whose predictive power for group sizes arises from the ability of households to deviate from any risk-sharing scheme jointly as 'sub-coalitions', as in Genicot and Ray (2003). Our estimation of the model showed that, for realistic income risk and preferences, this renegotiation-proof coalitional deviations model predicts insurance groups of up to five households, smaller than the village, and smaller also than typical exogenous groups such as extended families or castes within the village. Importantly, it is precisely the prediction of strong insurance in small insurance groups that enables the model to predict a realistic degree of insurance at the same time as symmetric responses of consumption to income shocks. Moreover, this is not a trivial feature of the model we propose: both in the standard limited commitment model and in a simple model of bufferstock savings consumption reacts substantially more to income rises than to income declines when the measured degree of insurance is in line with the data. These models can thus predict either the observed degree of insurance or approximate symmetry, but not both.

We think that our results are important not only for our understanding of risk sharing in rural India, but also for policymakers in developing countries more generally. This is because we would expect the effect of policy interventions such as poverty reduction or income insurance to change substantially once one allows for coalitional deviations from risk sharing. In more standard environments, where limited commitment arises from the possibility of individual deviations only, it has been shown that policies to reduce income risk may be counterproductive as they reduce the punishment of exclusion from insurance, thus making the outside option to the contract more attractive and reducing the transfers households are willing to make to others. This can potentially crowd out private consumption insurance (Attanasio and Rios-Rull, 2000; Krueger and Perri, 2011; Broer, 2011).[24] With endogenous group sizes, in contrast, crowding out works through a completely new channel by affecting the set of sustainable group sizes, which raises interesting questions for future research.

More generally, it would be interesting to perform a comprehensive comparative statics exercise that studies how group size, as well as the degree of insurance and symmetry of the resulting joint consumption-income process, respond to changes in the environment such as a change in income risk, access to formal financial markets, information frictions, and others.[25] Our estimate of small single-digit group sizes depends on the particular environment we consider, and the observed income and consumption characteristics in the ICRISAT data. Other contexts may imply much larger insurance groups. Finally, it would be interesting to analyse formally the dynamics governing the formation and break-up of risk-sharing groups when there are (anticipated) changes in the environment and groups are known to (potentially) be temporary, as in Bloch et al. (2007). We think that, beyond the application to village risk sharing used in the present study, these issues should also be important for analysing the stability of, for example, nation-states made up of different regions, or groups of countries that share risk in international organisations or regional unions.

24. Broer (2013) shows, however, that this is less likely in the context of developed countries, where individuals have more assets they can pledge under the contract, and income shocks are typically more persistent.

25. While previous studies have looked at limited information and limited commitment together (Broer et al., 2015), our focus on endogenous group formation makes the inclusion of any additional friction difficult.

References

Abreu, D., D. Pearce, and E. Stacchetti, "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica*, 1990, *58*, 1041–1063.

Ambrus, Attila, Markus Mobius, and Adam Szeidl, "Consumption Risk-Sharing in Social Networks," *American Economic Review*, 2014, *104* (1), 149–82.

Angelucci, Manuela, Giacomo de Giorgi, and Imran Rasul, "Resource Pooling within Family Networks: Insurance and Investment," 2015. Mimeo.
ΩÁbrahám and Laczó

Árpád Ábrahám and Sarolta Laczó, "Efficient Risk Sharing with Limited Commitment and Storage," *Review of Economic Studies*, 2018, *85* (3), 1389–1424.

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

Banerjee, Abhijit, Arun G. Chandrasekhar, Esther Duflo, and Matthew O. Jackson, "The Diffusion of Microfinance," *Science*, 2013, *341* (6144).

Bernheim, B. D. and D. Ray, "Collective Dynamic Consistency in Repeated Games," *Games and Economic Behavior*, 1989, *1*, 295–326.

_ , B. Peleg, and M. Whinston, "Coalition Proof Nash Equilibria. I: Concepts," *Journal of Economic Theory*, 1987, *42* (1), 1–12.

Binswanger, Hans P. and N.S. Jodha, "Manual of Instructions for Economic Investigators in ICRISAT's Village Level Studies," Technical Report, Village Level Studies Series, Economics Program, ICRISAT 1978.

Bloch, Francis, Garance Genicot, and Debraj Ray, "Reciprocity in Groups and the Limits to Social Capital," *American Economic Review*, 2007, *97* (2), 65–69.

_ , _ , and _ , "Informal insurance in social networks," *Journal of Economic Theory*, November 2008, *143* (1), 36–58.

Bold, Tessa, "Implications of endogenous group formation for efficient risk-sharing," *Economic Journal*, 2009, *119*, 562–591.

_ and Stefan Dercon, "Insurance Companies of the Poor," *CEPR Discussion Papers*, 2014, *10278.*

— **and Tobias Broer**, "A Note on Group Formation," December 2018. .

Broer, Tobias, "Crowding out and crowding in: When does redistribution improve risk-sharing in limited commitment economies?," *Journal of Economic Theory*, May 2011, *146* (3), 957–975.

— , "The Wrong Shape of Insurance? What Cross-Sectional Distributions Tell Us about Models of Consumption Smoothing," *American Economic Journal: Macroeconomics*, 2013, *5* (4), 107–40.

— **, Marek Kapička, and Paul Klein**, "Consumption Risk Sharing with Private Information and Limited Enforcement," CERGE-EI Working Papers wp531, The Center for Economic Research and Graduate Education - Economics Institute, Prague February 2015.

Chandrasekhar, A., C. Kinnan, and H. Larreguy, "Does Savings Crowd Out Informal Insurance? Evidence from a lab experiment in the field," Technical Report, MIT 2010.

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

Deaton, Angus, "On Risk, Insurance and Intra-Village Consumption Smoothing," 1990. Mimeo.

De Weerdt, J., *Insurance Against Poverty*, Oxford University Press, 2004.

Dubois, Pierre, "Heterogeneity of Preferences, Limited Commitment and Coalitions: Empirical Evidence on the Limits to Risk Sharing in Rural Pakistan," CEPR Discussion Papers 6004, C.E.P.R. Discussion Papers December 2006.

— **, Bruno Jullien, and Thierry Magnac**, "Formal and Informal Risk Sharing in LDCs: Theory and Empirical Evidence," *Econometrica*, 07 2008, *76* (4), 679–725.

Fafchamps, M. and S. Lund, "Risk-Sharing Networks in Rural Philippines," *Journal of Development Economics*, 2003, *71*, 261–287.

Farrell, J. and E. Maskin, "Renegotiation in Repeated Games," *Games and Economic Behavior*, 1989, *1*, 327–360.

Fitzsimons, Emla, Bansi Malde, and Marcos Vera-Hernandez, "Group Size and the Efficiency of Informal Risk Sharing," mimeo, University College London 2015.

Genicot, G. and D. Ray, "Group Formation in Risk-Sharing Arrangements," *Review of Economic Studies*, 2003, *70*, 87–113.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante, "Consumption and Labor Supply with Partial Insurance: An Analytical Framework," *American Economic Review*, July 2014, *104* (7), 2075–2126.

Kinnan, Cynthia, "Distinguishing barriers to insurance in Thai villages," 2014.

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

Krueger, Dirk and Fabrizio Perri, "Understanding Consumption Smoothing: Evidence from the U.S. Consumer Expenditure Data," *Journal of the European Economic Association*, 2005, *3* (2-3), 340–349.

_ and _ , "Public versus private risk sharing," *Journal of Economic Theory*, 2011, *146* (3), 920–956.

Laczó, Sarolta, "Risk Sharing With Limited Commitment and Preference Heterogeneity: Structural Estimation and Testing," *Journal of the European Economic Association*, 2014.

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

_ , J. P. Thomas, and T. Worrall, "Mutual Insurance, Individual Savings and Limited Commitment," *Review of Economic Dynamics*, 2000, *3*, 216–246.

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

Ligon, Ethan, *Insurance Against Poverty*, ed., Vol. of , : Oxford University Press, .

Lim, Youngjae and Robert M. Townsend, "General Equilibiurm Models of Financial Systems: Theory and Measurement in Village Economies," *Review of Economic Dynamics*, 1998, *1*, 59–118.

Mazzocco, Maurizio and Shiv Saini, "Testing Efficient Risk Sharing with Heterogeneous Risk Preferences," *American Economic Review*, February 2012, *102* (1), 428–68.

Meghir, Costas, Ahmed Mushfiq Mobarak, Corina D Mommaerts, and Melanie Morten, "Migration and Informal Insurance: Evidence from a Randomized Controlled Trial and a Structural Model," Working Paper 26082, National Bureau of Economic Research July 2019.

Mobarak, A. Mushfiq and Mark Rosenzweig, "Selling Formal Insurance to the Informally Insured," Technical Report, Yale University 2012.

Morduch, J., "Consumption Smoothing Across Space: Tests for Village-Level Responses to Risk," 1991. Mimeo, Harvard University.

Morten, Melanie, "Temporary migration and endogenous risk sharing in village india," Technical Report, National Bureau of Economic Research 2016.

Murgai, R., P. Winters, E. Sadoulet, and A. de Janvry, "Localized and Incomplete Mutual Insurance," *Journal of Development Economics*, 2002, *67*, 245–274.

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

Park, Yena, George Mailath, Dirk Krueger, and Harold Cole, "Fragile Financial Coalitions Under Belief Coordination Frictions," mimeo 2018.

Ravallion, M. and S. Chaudhuri, "Risk and Insurance in Village India: Comment," *Econometrica*, 1997, *65*, 171–184.

Ray, Debraj and Rajiv Vohra, "Coalition formation," in "Handbook of game theory with economic applications," Vol. 4, Elsevier, 2015, pp. 239–326.

Rouwenhorst, K. Geert, "Asset pricing implications of equilibrium business cycle models," in Thomas F. Cooley, ed., *Frontiers of business cycle research*, Princeton University Press, 1995, p. 294330.

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

Winter, Joachim K, Kathrin Schlafmann, and Ralf Rodepeter, "Rules of Thumb in Life-cycle Saving Decisions," *The Economic Journal*, 2012, *122* (560), 479–501.

Appendix: FOR ONLINE PUBLICATION

In this online appendix, we report auxiliary estimation results, such as the transition matrices for income we use in our computation (Section A.1) and additional descriptive statistics for the ICRISAT data (Section A.2). In Section A.3, we explain the simulated method of moments estimation procedure and in Section A.4, we illustrate the identification problem in the limited commitment model at high levels of insurance.

The remainder of the online appendix discusses the results and their robustness along a number of dimensions. In Section A.5, we evaluate the quality of the approximation of the CD and ID models that we use in our benchmark results relative to exact solutions of these models. Section A.6 studies whether there is a simpler rule-of-thumb implementation of the constrained-efficient contract that keeps the main features of the full model but is easier to implement.

Sections A.7–A.15 conduct a variety of robustness checks on the benchmark estimation. First, sections A.7–A.9 examine robustness of the finding of small insurance groups in the CD model. Section A.7 considers alternative outside options in the CD model. In Section A.8, we allow for households to have heterogeneous income processes in the CD model. Section A.9 examines sensitivity of the main results with respect to the assumption that deviating coalitions start out from equal sharing.

Next, sections A.10–A.13 examine the robustness of the large asymmetry in consumption and income co-movements in the ID model. In A.10, we use a different number of support points in the process of individual incomes. In A.11 both simulated and data moments are constructed without time-demeaning. A.12 discusses the role of income persistence and presents results for the CD and ID models where we also estimate income persistence to target the consumption moments. A.13 considers a simple form of preference heterogeneity.

Finally, we examine the robustness of the estimation results in the CD, ID and SI models with respect to how we treat unmodelled variation in the data. A.14 estimates the ICRISAT target moments and income process without controlling for household fixed effects. A.15 generalizes our benchmark estimates to include measurement error in consumption and incomes.

Table A.1. Transition probabilities for household incomes

	Aurepalle			**Kanzara**			**Shirapur**		
	y_1	y_2	y_3	y_1	y_2	y_3	y_1	y_2	y_3
y_1	0.4085	0.4613	0.1302	0.2487	0.5000	0.2513	0.1686	0.4840	0.3474
y_2	0.2306	0.5387	0.2306	0.2500	0.5000	0.2500	0.2420	0.5160	0.2420
y_3	0.1302	0.4613	0.4085	0.2513	0.5000	0.2487	0.3474	0.4840	0.1686

A.1. Transition matrices

Table 19 reports the probabilities of moving from income state y_i to y_j for Aurepalle, Kanzara, and Shirapur.

A.2. Descriptive statistics

Table A.2 presents descriptive statistics from the ICRISAT data.

Table A.2. Descriptive statistics

	Aurepalle		Kanzara		Shirapur	
Variable	Mean	Sd	Mean	Sd	Mean	Sd
Consumption	1622.10	704.31	2095.43	1113.91	2359.87	1075.85
Consumption (aeq.)	303.47	127.86	400.84	161.42	430.37	170.71
Income	3787.41	3734.31	5623.42	5524.55	4432.26	3490.73
Income (aeq.)	629.58	429.78	984.42	742.54	792.16	577.58
Aeq. household size	5.95	2.70	5.66	2.68	5.85	2.52
No. of observations	204		222		186	
No. of households	34		37		31	

Note: Monthly consumption and income measured in 1975 Indian rupees per year. In 1975, 8 Indian rupees were worth about 1 US dollar, which is about 4.60 dollars in 2016 (see Laczó (2014) for calculations).

A.3. Simulated method of moments

With the estimated income process and log-preferences as input, we solve the CD and comparison models on a fine discrete grid of $\delta \in [0.5, 0.99]$. We use the approximation procedure laid out in Section 2.3 and compute a numerical solution for the CD insurance contract in the largest stable group and for the ID insurance contract in the village. For the SI model, we compute the individual policy function.

We then draw a panel of income realisations for N households in the ID and SI models, and N' households in the CD model, over $T = 6,200$ periods (starting with savings equal to 0 in the case of the SI model). We then use the solutions to the ID and CD models, and the individual policy function in the SI model, to simulate a panel of individual consumption. After discarding the first 200 periods and subtracting period-specific village-averages from the simulated individual data (to make it comparable to our treatment of the ICRISAT data), we calculate the four moments from this simulated sample.

In order to estimate the discount factor in the ID, CD and SI model, the criterion to be minimised is

$$\Lambda(\delta) = [f - g(\delta)]' W^{-1} [f - g(\delta)] \tag{A.1}$$

where f is the vector of moments calculated from the ICRISAT data and $g(\delta)$ is the vector of simulated moments. That is, we minimize the weighted distance between the data moments and the simulated model moments. The weighting matrix W has on the diagonal the variances of the data moments, which are obtained by bootstrapping the data 1,000 times.

Importantly, with the simulated method of moments estimation we can also compare the difference between simulated and target moments and the minimized criterion function across the three models – and thus judge the empirical performance of the CD model relative to the alternatives.

A.4. Identification of preference parameters and borrowing constraints

As noted in Laczó (2014), a necessary condition for separate identification of the discount factor and the coefficient of risk aversion in the standard model with individual deviations is that households have binding constraints in as many income states as there are parameters to estimate. But because insurance is relatively strong in the ICRISAT villages, usually at most one constraint is binding. It is therefore not possible to identify time and risk preferences separately in either the ID or the CD model.[26] We therefore normalise the coefficient of relative risk aversion σ to 1 (log-preferences) in the estimation, and estimate the discount factor via the simulated method of moments.

Similarly, in the SI model, we cannot separately identify the discount factor and the interest rate and borrowing limit. We thus choose an interest rate equal to 4 percent and set the borrowing limit to 0. The reason for this choice is pragmatic: as it turns out, the moments we consider are not very sensitive to the choice of these parameters. With the interest rate and borrowing limit calibrated, we estimate the discount factor via the simulated method of moments.

Figure A.1 shows how neither the CD model (here depicted in its version with one period autarky following a deviation, as in Genicot and Ray (2003)) nor the ID model separately identify the two preference parameters, risk aversion σ and discount factor δ. Rather when σ declines – at a higher level in the case of coalitional deviations where average insurance is lower – as δ rises, neither the moments nor the goodness of fit changes. We thus normalise σ to 1 (log-preferenes) in the structural estimation. Note that the endogeneity of group size in the CD model can potentially help identification: at a discount factor above 0.96 we choose a risk aversion parameter that is consistent with unchanged moments and goodness of fit at a given group size of 4. But as the figure shows, now groups of 5 become sustainable, which increases, for example, the degree of insurance, and decreases the goodness of fit. This may bound the range of discount factors for which the model is not identified. Within those bounds, however, the model is still not identified. This is why we choose to normalize $\sigma = 1$ in both the ID and CD model.

A.5. Comparing the exact model and approximation

The analysis in our paper adapts the standard approximation of the dynamic limited commitment model (Ligon et al., 2002) that derives the policies for the Lagrange multipliers, φ, in equation (7) from a ficticious setting where an individual interacts with a homogeneous rest of the village whose outside

26. We compute the model for three individual income states, which is the minimum needed for identification. However, increasing the number of income states does not yield identification either, simply because the targeted degree of insurance is too high.

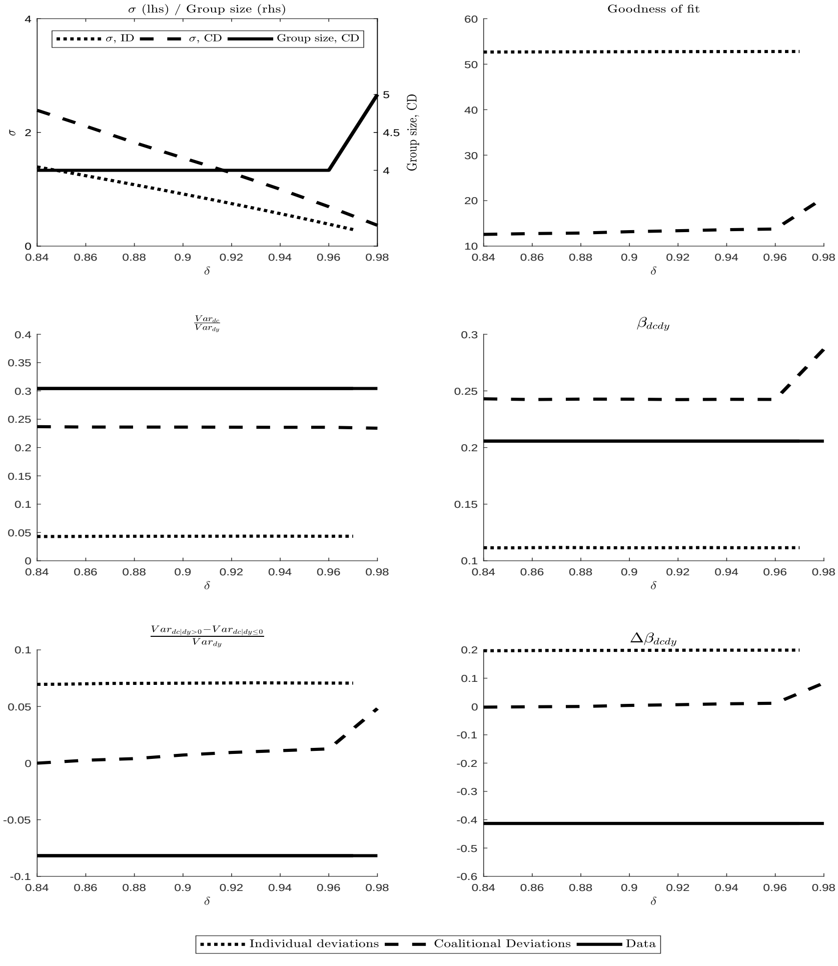

FIGURE A.1. Estimated parameters and moments in Aurepalle

The figure shows the estimated risk aversion σ and (in the case of coalitional deviations) group size (panel 1), the model criterion (panel 2) and the 4 moments of interest for different values of the discount factor δ.

option depends only on aggregate income. The approximation then calculates consumption shares using the relative multipliers that result when several such individuals interact. While this approximation is exact for $n = 2$ and accurate for large groups of several hundred households (Ligon et al., 2002), how it evolves for $n > 2$ is unknown. This appendix compares the approximation to an exact solution along two dimensions: first, we compute the exact solution of both the CD and ID models for small groups of size 3 and 4, the largest sizes where this is feasible, and compare the implied features to those of the approximation. Second, we identify the maximum sustainable size of insurance groups exactly, under the assumption of full insurance.

A.5.1. Exact and approximate model solutions at small group sizes . While the approximation trivially coincides with the exact solution when insurance is perfect, we expect it to be imperfect whenever risk sharing is partial. In particular, the assumption of income-pooling in the rest of the village in the approximation introduces two sources of error relative to the exact solution: (i) the Lagrange multiplier on the incentive constraint of the rest of the village is calculated at its average income realisation, but in an exact solution, it is the average of the Lagrange multipliers across the income distribution in the rest of the village that determines consumption of any remaining unconstrained individual. Given the convexity of the Lagrange multiplier around zero, the multiplier at the average income realisation will in general be smaller than the average of the multipliers across the idiosyncratic income realisations. Hence, for given preferences the approximation will tend to predict more risk sharing than the exact solution and this difference increases with group size. (ii) in an exact solution, deviating individuals continue in autarky (in the ID model) or in a smaller stable group (in the CD model). This contrasts with an outside option of perfect risk sharing within the rest-of-the-village in the approximation. Without persistence in incomes (as is approximately true in our data), this makes the outside option strictly better in the approximation and hence its participation constraint more binding, resulting in less risk sharing (and thus potentially offsetting effect (i)). In the CD model, there is an additional source of error resulting from our assumption of equal treatment (subject to binding participation constraints) following a deviation. This prevents the planner from treating coalition members asymmetrically despite the fact that this might increase her ability to deter joint deviations and may be optimal given the history of the contract. Importantly, while it is a priori unclear which of the three

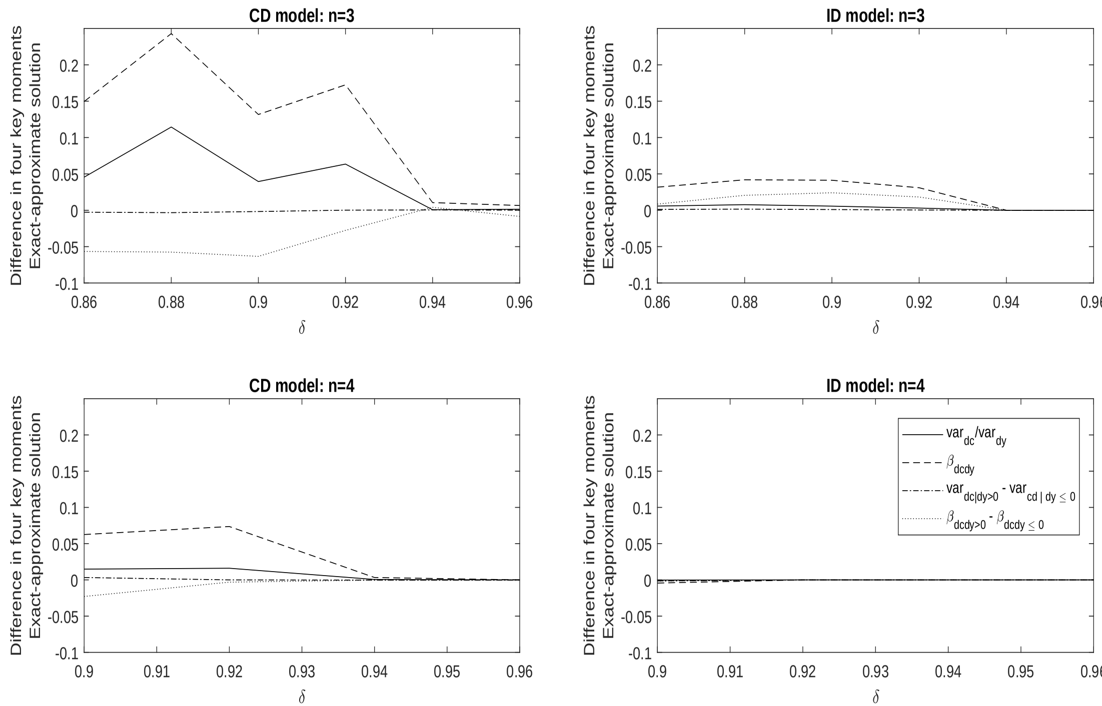

FIGURE A.2. Simulated moments from exact solution of CD and ID model for groups of size three and four.

The table shows four key moments summarising the joint distribution of consumption and income growth in simulations of the CD model vs. the ID model with groups of size 3 and 4 and $\delta \in [0.86, 0.96]$. In each panel, the solid black line plots the difference between exact and approximate solution in the ratio of the variance of individual consumption changes to the variance of individual income changes. The dashed line plots the difference between exact and approximate solution in the slope coefficient in a regression of consumption growth on income growth. The dashed-dotted line plots the difference between exact and approximate solution in the asymmetry of the variance ratio for income winners and losers. The dotted line reports the difference between exact and approximate solution in the asymmetry of the slope coefficients in a regression of consumption growth on income growth.

effects dominates, approximation error should decrease the higher is the degree of insurance, and the less often participation constraints bind in equilibrium.

We now compare the exact solution, based on solving equations (16) – (19), and the approximate solution for the small group sizes, $n = 3$ and $n = 4$, where we can calculate the former. Note that, while Bold (2009) performs a similar analysis of the model in groups of three without income persistence, we are, to our knowledge, the first to study the exact solution of the three and four agent-limited commitment model with an estimated income process and compare it to data from actual village economies.

In Figure A.2, we compare the approximate and the exact solution of ID and CD models. We solve the models with log utility and discount factors δ that generate moderate to high insurance, for the income process estimated in Aurepalle.[27,28] The figure plots the difference between the exact and approximate versions of the four key moments used in the analysis, namely measures of the sensitivity of consumption with respect to income changes and its asymmetry in the sub-samples of income winners and losers.

The approximation errors in the CD model are indeed very small at high values of δ similar to the ones we predict in the benchmark estimation in Table 5, giving us confidence in our main result. And the same holds for the small-group version of the ID model. Again, this is because any difference between the exact and approximate solutions arises from states when participation constraints bind, and there are fewer of those at high discount factors.

This also implies that we would expect approximation errors to rise at lower discount factors. Figure A.2 shows this to be the case for both the ID and CD models at $n = 3$, if much less so for $n = 4$, where insurance is higher for a given discount factor in the ID model and CD groups are not stable for discount factors below .9. Specifically, the exact solution of the CD model predicts somewhat lower

27. We use an income process whose variance is estimated using income data from Aurepalle, but with lower persistence, which makes a numerical solution of the exact model easier.

28. For discount factors lower than .94, groups are not stable in either the exact or the approximate solution. To make comparisons over a meaningful range, we thus introduce a state-independent penalty, that is chosen to be as small as possible (up to a maximum of 20% of the outside option) such that the group size under consideration is stable.

insurance than the approximation at low discount factors, suggesting that the use of an average participation constraint for the rest of village in the approximate solution (as discussed above) may be the dominant source of error.

Overall, the increase in the approximation error at lower discount factors seems modest in relation to the level of the moments (see Figure A.2), although, as one might expect, the approximate solution does understate somewhat the increasing difference in the predicted degree of insurance between the CD and ID models in groups of 3 households. Importantly, the rise in approximation errors is concentrated at group sizes and values of δ smaller than those we estimate. Our conclusion from this analysis is therefore that for the income processes and high degrees of insurance observed in the ICRISAT data, and the resulting estimated group sizes, the approximation we use is accurate.

A.5.2. The set of stable group sizes with full insurance. The results in the previous subsection make us confident that our results accurately capture the features implied by limited commitment insurance in small groups. Since we cannot solve the models exactly for larger groups, however, these results are strictly speaking informative only under the maintained assumption that the approximation accurately identifies stable groups. That is, our comparison of the exact and approximate solution at small group sizes justifies the use of the approximation only inasmuch as there are no larger group sizes that would be stable in the exact (but not the approximate) solution. In this section, we examine this assumption.

While we cannot solve the exact model in its general version for larger group sizes, we can identify sustainable groups in the exact model whenever perfect insurance is sustainable. This is relevant in particular as insurance is indeed close to perfect in our benchmark estimation. In this section, we therefore look at sustainable groups with perfect income-pooling. Specifically, we maintain equal sharing of consumption resources in groups whose size varies between $n = 2$ and the village size, and identify the maximum group size such that individuals would in no income state find it optimal to deviate and form any smaller group (that continues to share resources equally after one period of autarky).

To understand the results in Table A.3, note that with full insurance, the predicted moments are insensitive to changes in preferences that are consistent with a given sustainable group size. In other words, even after normalizing the risk aversion parameter, the discount factor δ is only set identified between a lower and an upper value, δ_{min} and δ_{max} respectively, indicated in the table. It turns out that this range only comprises one value of δ on our grid, however, which is, as expected, higher than in our benchmark estimates, as required to make full insurance sustainable relative to the outside option of individual autarky. Importantly, the estimated maximum sustainable group sizes are identical to our benchmark estimates. Again, these results give us confidence that our benchmark estimation, based on our approximation, accurately captures the data generating process.

TABLE A.3. Sustainable groups with full insurance

	Aurepalle		**Kanzara**		**Shirapur**	
	Data	**CD**	**Data**	**CD**	**Data**	**CD**
n		4.00		4.00		4.00
δ_{min}		0.97		0.98		0.97
δ_{max}		0.98		0.98		0.97
σ		1.00		1.00		1.00
$\frac{Var_{dc}}{Var_{dy}}$	0.30	0.23	0.56	0.24	0.33	0.23
β_{dcdy}	0.21	0.23	0.22	0.23	0.17	0.22
$\frac{Var_{dc\|dy>0}-Var_{dc\|dy\leq 0}}{Var_{dy}}$	-0.08	0.00	-0.25	0.00	-0.16	0.00
$\beta_{dcdy>0} - \beta_{dcdy\leq 0}$	-0.41	-0.00	-0.14	0.00	-0.05	-0.00
Model criterion		12.28		5.07		6.14

Note: For each of the three ICRISAT villages, the table shows four key moments summarising the joint distribution of consumption and income growth in the survey ("Data", in the first column for each village), and in simulations where groups fully insure consumption of its members, but are required to be sustainable in the long run with respect to deviations to smaller full-insurance groups. For the simulated model solutions, the table also presents the size of the insurance groups n and the discount factor δ, and measurement error in consumption and income $Var_{d\varepsilon}/Var_{dc}$ and $Var_{d\nu}/Var_{dy}$. δ, $Var_{d\varepsilon}/Var_{dc}$ and $Var_{d\nu}/Var_{dy}$ are chosen to minimise the sum of differences between all four moments predicted by the models and those observed in the data, weighted by the inverse variance of the data moments calculated using a bootstrap procedure. The estimated parameters are those that minimise this criterion on a grid of $\delta \in [0.9, 0.99]$. The last line reports the value of the criterion at the estimated parameters.

A.6. Implementing the constrained-optimal contract with a rule of thumb

The transfer function in the CD model is potentially complex. In this subsection, we examine whether there is a simpler rule of thumb for the constrained-efficient contract that fulfills the following properties (see Winter et al. (2012) for a related discussion in the context of the optimal savings problem): (i) the rule of thumb for transfers is simple and closed-form, (ii) it retains the essence of the exact solution, namely history dependence (whereby households are rewarded in the future for contributions they make today) and coalition-proofness.

In general, a closed-form solution of the limited commitment risk-sharing contract can only be obtained when a candidate transfer rule is determined independently of participation constraints. To find a rule of thumb, we therefore do not look for the constrained optimal transfer that satisfies the participation constraints exactly (as in Section 2), but rather examine whether a given fixed transfer rule is stable with respect to deviations. The simplest fixed transfer rule is presumably equal sharing of resources among households. Inspired by this, we consider a simple setting, in which households are either classed as high-income or low-income (i.e. we consider only two income states) and concentrate on a class of simple transfer rules according to which high-income households transfer a constant percentage x of the equal sharing transfer to low-income households every period. To this, we add a simple form of history dependence: when all group members have the same income, those who have received transfers in the previous period repay y% of the received transfer to those who made it.[29]

The set of stable groups and insurance contracts is then derived recursively, as in the full model: Suppose that for each $m < n$, we have identified the transfer rule $\sigma(x^*, y^*, m)$ that maximizes expected discounted life-time utility subject to the insurance contract being stable with respect to deviations by sub-groups (which themselves must be stable with respect to further deviations). For a group of size n, we then calculate for each insurance contract $\sigma(x, y, n)$ with $x \in (50, 100)$ and $y \in (0, 25)$ the implied discounted life-time utility. We then check in each state whether the utility the k individuals with a high income realization derive from the contract is higher than their expected payoff of deviating as a group

29. We restrict history dependence to these symmetric states because this is when the full model suggests that history dependence matters most: participation constraints typically do not bind in these states, such that consumption shares are a direct function of the history of states and transfers through the updated relative Pareto weights, γ_r^i for $i = 1, ...n-1$.

with $m \leq k$ individuals that implements its preferred stable contract $\sigma(x^*, y^*, m)$ following deviation (since this is the most binding constraint).

If all participation constraints are satisfied for $\sigma(x, y, n)$, then this contract is part of the set of stable contracts for group n. Having made this calculation for each possible contract $\sigma(x, y, n)$, the constrained-optimal contract is that one among the stable ones, which maximizes expected discounted life time utility. If no contract satisfies the participation constraints, the group of size n is deemed unstable.

Implementing this rule of thumb for the CD model, we estimate group sizes of 4 in all villages, very similar to the full model. Rather than close-to perfect insurance in small groups, as in the full model, the rule of thumb predicts full risk sharing in these groups ($x = 100$, $y = 0$). In other words, the rule-of-thumb estimates are equivalent to those in Section A.5.2, where we looked at maximum stable group sizes under the maintained assumption of full insurance (and we therefore omit reporting them).[30] The simpler models match the data almost as well as the exact or more complex approximation of the contract. Hence, it is possible that the observed data is generated by a rule-of-thumb implementation of the constrained-efficient risk-sharing contract that is robust to coalitional deviations.

A.7. Alternative outside options

In this section, we briefly investigate how the predictions of the CD model change when we consider alternative specifications of the outside option.

A.7.1. One period of autarky. In our benchmark setting of the CD model, we allowed households to share risk in a smaller coalition starting from the first period of the deviation. This section studies how the results are changed when we assume that individuals go through one period of autarky before they can renegotiate and share risk with others (as in Genicot and Ray (2003)).

30. Our quantitative analysis of rules of thumb only considers two income states which leads to small quantitative differences in the moments relative to those reported in Table A.3.

Table A.4 shows how the effect of this less attractive outside option (that acts to increase insurance) is mainly offset by a lower estimated discount factor, leaving the predicted moments close to their benchmark values. The maximum group size is unchanged in Aurepalle and Kanzara, but increased to 5 in Shirapur. The asymmetry continues to be close to zero and the goodness of fit is only marginally inferior to that of our benchmark estimates.

TABLE A.4. Preferences estimated to target all 4 moments - one period autarky

	Aurepalle					
	Data	**CD**	**Data**	**CD**	**Data**	**CD**
n		4.00		4.00		5.00
δ		0.94		0.94		0.92
s.e.		0.01		0.01		0.01
σ		1.00		1.00		1.00
$\frac{Var_{dc}}{Var_{dy}}$	0.30	0.24	0.56	0.25	0.33	0.18
β_{dcdy}	0.21	0.24	0.22	0.28	0.17	0.21
$\frac{Var_{dc\|dy>0}-Var_{dc\|dy\leq 0}}{Var_{dy}}$	-0.08	0.01	-0.25	0.01	-0.16	0.01
$\beta_{dcdy>0} - \beta_{dcdy\leq 0}$	-0.41	0.00	-0.14	0.03	-0.05	0.02
Model criterion		12.99		5.92		8.37

Note: The table shows four key moments summarising the joint distribution of consumption and income growth in the survey ("Data", in the first column for each village), and in simulations of the CD model without a period of autarky following a deviation (in the second column). For the simulated model solutions, the table also presents the size of the insurance groups n and the discount factor δ, which is chosen to minimise the sum of differences between all four moments predicted by the models and those observed in the data, weighted by the inverse variance of the data moments calculated using a bootstrap procedure. The estimated preference parameters are those that minimise this criterion on a grid of $\delta \in [0.5, 0.99]$. The last line reports the value of the criterion at the estimated parameters.

A.7.2. Second-best risk sharing in the 'rest-of-village'. The approximate solution of the CD model follows Ligon et al. (2002) by studying a two-agent contract between an individual and the rest of an insurance group, and by approximating the rest of the group as a single agent in the solution of the constrained-efficient allocation (although not in the simulation, where the full vector of incomes is taken into account). This simplification corresponds to assuming, for the solution of the contract, that the rest of the village is able to share risk perfectly. This implies an inconsistency, as the households in the rest of the village are better able to share risk among themselves than with the individual under consideration. In principle, one could try to find a fixed point by iteratively adjusting the outside option for the rest of the village to be consistent with a candidate solution to the two-agent problem until the two imply the same degree of risk sharing. This is, unfortunately, computationally too demanding for our estimation approach. The recursive solution of the CD model, however, where we sequentially solve for the contract in groups of $n = 2, 3, 4$ etc., offers an alternative approximation of rest-of-group utility. Specifically, it allows us to use as the outside option of the rest of the group the average expected utility that its members expect to get from a risk-sharing contract with $n - 1$ households. Making the same assumption as in our benchmark approximation that deviating households share the surplus of the group as equally as possible from the moment of deviation then allows us to calculate an alternative approximate solution of contracts in the CD model where there are the same frictions within the rest of the village as between it and the individual under consideration. Intuitively, we would expect this outside option to improve risk sharing between the individual and the rest of the group, whose less attractive outside option of constrained (as opposed to perfect) risk sharing corresponds to a relaxation of its participation constraints relative to the benchmark model.

Table A.5 presents the results: they are, essentially, identical to our benchmark estimates, which is unsurprising given the high degree of insurance in all subgroups at the estimated discount factors.

A.8. More income heterogeneity

In our benchmark model in Section 2, we assume that households are ex-ante identical and differ ex-post only in their income realization. In this section, we allow households to differ ex-ante by considering heterogeneous income processes. Specifically, we follow Laczó (2014) and divide households in the village into two groups depending on whether a household's mean income is above or below the median village income over the six panel rounds. We then estimate the income process using the methods in Section 4.1.2 separately for the households with 'low' mean income and those with 'high' mean income. The estimation results are shown in Table A.6. Mean income for the 'low income' process is one third of the average income for the 'high income' process (335 rupees compared with 924 rupees). The 'low income' process also has a higher persistence and lower variance than the 'high income' process and is thus less risky. To discretize these income

TABLE A.5. Preferences estimated to target all 4 moments - second-best outside option for the rest of the village

	Aurepalle					
	Data	**CD**	**Data**	**CD**	**Data**	**CD**
n		4.00		4.00		4.00
δ		0.97		0.98		0.96
s.e.		0.00		0.00		0.01
σ		1.00		1.00		1.00
$\frac{Var_{dc}}{Var_{dy}}$	0.30	0.23	0.56	0.24	0.33	0.23
β_{dcdy}	0.21	0.23	0.22	0.23	0.17	0.22
$\frac{Var_{dc\|dy>0}-Var_{dc\|dy\leq 0}}{Var_{dy}}$	-0.08	0.00	-0.25	0.00	-0.16	0.00
$\beta_{dcdy>0}-\beta_{dcdy\leq 0}$	-0.41	-0.00	-0.14	0.00	-0.05	-0.00
Model criterion		12.28		5.07		6.14

Note: The table shows four key moments summarising the joint distribution of consumption and income growth in the survey ("Data", in the first column for each village), and in simulations of the CD model where the outside option of the the rest of the village in the approximation does not assume perfect risk sharing as detailed in the main text (in the second column for each village). For the simulated model solutions, the table also presents the size of the insurance groups n and the discount factor δ, which is chosen to minimise the sum of differences between all four moments predicted by the models and those observed in the data, weighted by the inverse variance of the data moments calculated using a bootstrap procedure. The estimated preference parameters are those that minimise this criterion on a grid of $\delta \in [0.5, 0.99]$. The last line reports the value of the criterion at the estimated parameters.

processes, we use the same method as in the benchmark model, normalizing the average income across both income processes to 1.

We solve this model for the village of Aurepalle on a reduced grid for the discount factor ($\delta = [0.9, 0.98]$) and considering a maximum stable group size of 7. To calculate the simulated moments from the model solutions when income processes are heterogeneous, we need to take a stand on which households form stable insurance groups with each other. As in the benchmark estimation, we assume that the village splits into k equal-sized groups. Second, we examine two sorting processes of individuals into these insurance groups (a complete analysis of the sorting process is beyond the scope of the paper): in the first scenario, all insurance groups are homogeneous (i.e. half of them have only low-income and half of them have only high-income members). In the second case, all insurance groups are heterogeneous, but symmetric in the sense that the village consists of $k/2$ groups with, say, one low-income and $n-1$ high-income process individuals,

Table A.6. Estimated income processes

	Aurepalle	
	Low mean income	High mean income
Mean income	335.16	923.99
ρ	0.43	0.19
Var_{α_i}	0.28	0.11
Var_{ε}	0.13	0.16

Note: Households in the village of Aurepalle are separated into two groups, those with below median income, and those with above median income. The table presents the point estimates for the persistence parameter ρ and the shock variance Var_{ε} for the AR(1) process (27), as well as the variance of household fixed effects Var_{α_i}, in each of these groups. Bootstrapped standard errors in parentheses.

and $k/2$ groups with one high-income and $n-1$ low income process individuals (since, by construction, the village consists of equal numbers of individuals with the two income processes).[31]

The results are shown in Table A.7. In column (1), we present the results for homogeneous groups and in column (2) for heterogenous groups.

For homogeneous group formation, the best fit is achieved with a discount factor of 0.98, where the largest stable group size is 4, very similar to the benchmark estimation.

When we allow insurance groups to consist of a mix of households with heterogeneous income processes, the estimated discount factor and group size decrease to 0.96 and 3, with groups consisting of one low-income and two-high income individuals. As a result, the goodness of fit is worse. The reason is that groups with a heterogeneous composition struggle to attain stability even for small groups. Specifically, groups that contain half or more low-income individuals are not stable.

Interestingly, the relative size of the variance moment and the sensitivity moment (though both too high) is more similar to the data. To understand the reason behind the increased relative variance of individual consumption growth, note that total income of heterogeneous groups is dominated by the

31. We allow for some flexibility in partitioning the N individuals. For example, we also allow for village with k groups of $\frac{n-1}{2}$ low-income individuals and $\frac{n+1}{2}$ high-income individuals.

income of households with high average income. Income asymmetry thus makes diversification through pooling less powerful. Since insurance is close to perfect in our estimation individual consumption moves closely in line with group income. Taken together, this implies that the variance of individual consumption growth increases in asymmetric groups. Since the variance of changes in individual log income does not depend on the composition of insurance groups (nor on average income levels), this results in an increase of the relative variance moment in heterogeneous insurance groups relative to the sensitivity moment, which is not seen in our benchmark results.

TABLE A.7. Preferences estimated to target all 4 moments - heterogeneous income process with heterogeneous and homogeneous insurance groups

	Aurepalle		
	Data	**CD**	
		(1) homogenous groups	(2) heterogenous groups
n		4.00	3.00
δ		0.98	0.96
σ		1.00	1.00
$\frac{Var_{dc}}{Var_{dy}}$	0.30	0.23	0.40
β_{dcdy}	0.21	0.23	0.31
$\frac{Var_{dc\|dy>0}-Var_{dc\|dy\leq 0}}{Var_{dy}}$	-0.08	-0.0002	0.0033
$\beta_{dcdy>0} - \beta_{dcdy\leq 0}$	-0.41	-0.0040	0.0015
Model criterion		11.84	17.27

Note: The table shows four key moments summarising the joint distribution of consumption and income growth in the survey ("Data", in the first column for each village), and in simulations of the CD model, which allow for income heterogeneity, namely a 'low' and a 'high' mean income process with idiosyncratic variations. In column (1), insurance groups are assumed to be homogeneous, so that only agents with the same income process form groups and the village consists of 50% groups with the 'low' and 50% groups with the 'high' income process. In column (2), the village is assumed to consist of heterogeneous insurance groups, where each insurance group contains of one agent with the low income process and two agents with the high income process. For the simulated model solutions, the table presents the size of the insurance groups n and the discount factor δ, which is chosen to minimise the sum of differences between all four moments predicted by the models and those observed in the data, weighted by the inverse variance of the data moments calculated using a bootstrap procedure. The estimated preference parameters are those that minimise this criterion on a grid of $\delta \in [0.9, 0.98]$ and allowing a maximum village size of 7. The last line reports the value of the criterion at the estimated parameters.

It is not possible, however, to implement this robustness check for large villages and all parameter combinations, as allowing for income heterogeneity increases the set of possible coalitional deviations to the power set of the village, with heterogenous deviation payoff vectors for each of its elements.

A.9. Deviators continue with existing division of surplus

In this section, we depart from the assumption that deviating subgroups start risk-sharing with the most equal division of surpluses. Instead, we assume that a deviating subgroup (consisting of the individual and a subset of the rest of the village) continue sharing risk with an initial division of surpluses inherited from the ambient group. To be precise, if the Lagrange multiplier on the individual's promise-keeping constraint in the existing group is 2, which, given log-utility, implies the individual would receive two thirds of pooled group income and the rest-of-the village receives one third, then a joint deviation of the individual and a subset of the rest-of-the village would also start with this 'history-dependent' sharing-rule (which may be adjusted immediately as enforcement constraints bind).

We estimate this model for the village of Aurepalle on a reduced grid for the discount factor ($\delta = [0.9, 0.98]$) and considering a maximum stable group size of 7. The results are shown in Table A.8 and are almost identical to the benchmark results, with a slightly improved fit.

Thus, we conclude that our results of a superior fit of the CD model are not dependent on imposing equal sharing after deviation. The reason that the model does not perform drastically different is that the CD enforcement constraints prevent very unequal treatment of the agent versus the rest of the village in equilibrium and the relative ratio of Pareto weights therefore never departs very far from the equal sharing that we impose in the main part of the paper.

It is not possible, however, to implement this robustness check for large villages and all parameter combinations. One of the reasons why we choose a deviation payoff that is independent of the previous history is that it allows us to solve the enforcement constraint simultaneously for all grid points on the grid of relative Pareto weights at which a particular coalition has a binding constraint relative to the first-best allocation in the contract. This is the case because once we have identified a coalition as having a binding outside option, we then set their payoff from staying inside the contract equal to this outside option, which only depends on the income realization, but not on the relative Pareto weights. This reduces computing time substantially. In contrast, allowing the outside option to also depend on the relative Pareto weights means we have to solve the enforcement constraint separately for each grid point at which a particular coalition is constrained.

TABLE A.8. Preferences estimated to target all 4 moments - deviators continue with the division of surplus in the ambient group

	Aurepalle	
	Data	**CD**
n		4.00
δ		0.97
s.e.		0.00
σ		1.00
$\frac{Var_{dc}}{Var_{dy}}$	0.30	0.23
β_{dcdy}	0.21	0.23
$\frac{Var_{dc\vert dy>0} - Var_{dc\vert dy\leq 0}}{Var_{dy}}$	-0.08	0.0004
$\beta_{dcdy>0} - \beta_{dcdy\leq 0}$	-0.41	-0.004
Model criterion		11.82

Note: The table shows four key moments summarising the joint distribution of consumption and income growth in the survey ("Data", in the first column for each village), and in simulations of the CD model where the outside option of the the rest of the village in the approximation does not assume perfect risk sharing as detailed in the main text (in the second column for each village). For the simulated model solutions, the table also presents the size of the insurance groups n and the discount factor δ, which is chosen to minimise the sum of differences between all four moments predicted by the models and those observed in the data, weighted by the inverse variance of the data moments calculated using a bootstrap procedure. The estimated preference parameters are those that minimise this criterion on a grid of $\delta \in [0.9, 0.98]$ and allowing a maximum village size of 7. The last line reports the value of the criterion at the estimated parameters.

Despite some limitations, the robustness checks conducted in Sections A.5–A.9 leave us confident that our finding of small insurance groups with high degrees of insurance in the CD model is not sensitive to the assumptions we make in the benchmark estimation. We now turn to examine the robustness of the finding of a large asymmetry in the ID model.

A.10. A finer income grid

For our simulations, we use a discrete version of the estimated income process in Table 3. Since most of our results focus on (moments of) the joint distribution of consumption and income growth, the way in which we discretize incomes might potentially affect our results. In particular, one might worry that the degree of asymmetry in the ID model, which was a primary reason for its inferior fit, may be a consequence of how we discretize the income process.

Although we are constrained in the number of income states for our benchmark results (which use some aspects of the joint distribution of incomes in the insurance group as a state variable), we can compute the ID model for different income processes using a procedure similar to that in Laczó (2014), where only the aggregate income in the rest of the village is relevant. Table A.9 reports the results for one such exercise, where we double the number of support points of individual income from 3 to 6 and approximate the income process for the rest of the village as a discretized AR(1) process with 5 support points.[32]

To understand the results, note that more dispersion in incomes (of either the individual or the rest of the village) reduces insurance. This is because higher maximum income typically implies a higher maximum autarky value, and therefore a more binding participation constraint for the highest income individuals. This explains the stronger insurance in this specification of the model (with a coarse approximation of rest-of-village income) relative to our benchmark: the former does not allow for the extreme income realizations for the rest of village that are present in our benchmark economy (with its full set of possible cross-sectional distributions in the rest of village).

When the process for individual income has three support points ($n_y = 3$), insurance is strong, and in fact close to perfect for the village of Shirapur (where the negative serial correlation in incomes implies less dispersed outside options). In line with the intuition in the previous paragraph, increasing the support points of the individual income process from three to six implies more extreme individual income realizations, and thus more extreme values of autarky, more binding participation constraints, and therefore less insurance. Importantly, however,

32. As in Laczó (2014), we estimate a continuous AR(1) process on simulated data and use the Rouwenhorst (1995) method to dicretize it.

the asymmetry of the standard model increases as the degree of insurance falls. Thus, whenever there is strong insurance with a coarse income support (as in our benchmark estimates), an increase in the income support lowers the degree of insurance but increases the asymmetry.

TABLE A.9. The ID model with different income processes

	Aurepalle			**Kanzara**			**Shirapur**		
	Data	**ID, n_y=3**	**ID, n_y=6**	**Data**	**ID, n_y=3**	**ID, n_y=6**	**Data**	**ID, n_y=3**	**ID, n_y=6**
n		34.00	34.00		37.00	37.00		31.00	31.00
β		0.89	0.89		0.90	0.90		0.88	0.88
$\frac{Var_{dc}}{Var_{dy}}$	0.30	0.04	0.05	0.56	0.06	0.08	0.33	0.00	0.01
β_{dcdy}	0.21	0.11	0.11	0.22	0.17	0.18	0.17	0.02	0.05
$\frac{Var_{dc\|dy>0}}{Var_{dc\|dy\leq 0}}$	-0.08	0.06	0.07	-0.25	0.09	0.11	-0.16	0.00	0.02
$\frac{\beta dcdy>0}{\beta dcdy\leq 0}$	-0.41	0.19	0.20	-0.14	0.24	0.28	-0.05	0.03	0.11

Note: For each of the three ICRISAT villages, the table shows four key moments summarising the joint distribution of consumption and income growth in the survey ("Data", in the first column for each village), and in simulations of the ID model with two discretizations of the individual income process (27) that both use the Rouwenhorst (1995) method but differ in the number of support points n_y, equal to three and six, respectively. The income process for the rest of the village is approximated as a discretized AR(1) process with 5 support points.

A.11. Estimating the model without time-demeaning

It is standard practice to evaluate the performance of risk sharing models by conditioning consumption and income data both from the ICRISAT panel and from model simulations on movements in aggregate resources, using residuals from a regression on time dummies (or by demeaning the sample period-by-period). As we discussed in Section 4.3, this procedure has somewhat different effects in the two models we analyse. In the standard model, with individual deviations, the only risk-sharing group coincides with the village. Conditioning on village-level aggregate income (equal to aggregate consumption) thus isolates the idiosyncratic movements in income and consumption. The alternative model, with coalitional deviations, however, predicted a village to consist of several insurance groups. Demeaning the sample period-by-period, therefore, does not eliminate fluctuations in group-level incomes, but only in village-level incomes. Since the remaining fluctuations in group-level income are symmetric and translate to individual consumption fluctuations, this may increase the symmetry in the alternative model. This section therefore presents estimates of the CD, ID and SI models without time-demeaning the data and simulated moments.

TABLE A.10. Estimated income process without time-demeaning

	Aurepalle	**Kanzara**	**Shirapur**
ρ	0.20	0.039	-0.17
Var_{α_i}	0.29	0.27	0.36
Var_{ε}	0.19	0.078	0.12

Note: The table presents the point estimates for the persistence parameter ρ and the shock variance Var_{ε} for the AR(1) process (27) without time-demeaning.

Table A.10 reports estimates of the income process (27) without time-demeaning, which mechanically increases the variance of shocks but has an ambiguous effect on the AR(1) parameter ρ. The estimated process is virtually identical to the benchmark in Shirapur, but slightly less (more) persistent in Aurepalle (Kanzara). The increase in the variance of shocks is most pronounced in Aurepalle.

In Table A.11 we repeat our benchmark estimation based on this data. This makes little difference to the degree of insurance measured in the data for the three villages, but reduces the negative asymmetry observed in all three villages.

The group sizes estimated in the CD model are unchanged. Its predicted degree of insurance and the predicted symmetry, are essentially identical to the benchmark estimation. The predictions of the two comparison models change only very little and the prediction of too much insurance and asymmetry in the opposite direction as the data remain.

Table A.11. Preferences estimated to target all 4 moments without time-demeaning

	Aurepalle				Kanzara				Shirapur			
	Data	**CD**	**ID**	**SI**	**Data**	**CD**	**ID**	**SI**	**Data**	**CD**	**ID**	**SI**
n		4.00	34.00			4.00	37.00			4.00	31.00	
δ		0.97	0.88	0.92		0.98	0.90	0.95		0.96	0.88	0.94
s.e.		0.01	0.01	0.01		0.01	0.01	0.01		0.01	0.02	0.01
σ		1.00	1.00	1.00		1.00	1.00	1.00		1.00	1.00	1.00
$\frac{Var_{dc}}{Var_{dy}}$	0.33	0.23	0.02	0.12	0.61	0.24	0.05	0.07	0.37	0.23	0.02	0.06
β_{dcdy}	0.22	0.23	0.08	0.26	0.18	0.23	0.14	0.17	0.15	0.22	0.07	0.16
$\frac{Var_{dc\mid dy>0}-Var_{dc\mid dy\leq 0}}{Var_{dy}}$	-0.01	-0.00	0.04	0.07	-0.17	0.00	0.08	0.04	-0.14	0.00	0.03	0.04
$\beta_{dcdy>0}-\beta_{dcdy\leq 0}$	-0.39	-0.00	0.15	0.22	-0.02	0.01	0.24	0.16	0.01	-0.00	0.13	0.17
Model criterion		15.83	60.12	46.93		5.50	13.33	10.77		5.43	20.05	16.30

Note: For each of the three ICRISAT villages, the table shows four key moments summarising the joint distribution of consumption and income growth in the raw data of the survey ("Data", in the first column for each village), and in simulations using the income process estimated with unconditional data (see Table A.10) in the CD model (in the second column for each village), the ID model (third column) and the SI model (fourth column). For the simulated model solutions, the table also presents the size of the insurance groups n (in the case of the ID and CD models) and the discount factor δ, which is chosen to minimise the sum of differences between all four moments predicted by the models and those observed in the data, weighted by the inverse variance of the data moments calculated using a bootstrap procedure. The estimated preference parameters are those that minimise this criterion on a grid of $\delta \in [0.5, 0.99]$. The last line reports the value of the criterion at the estimated parameters.

A.12. The role of income persistence

The persistence of income shocks is a key parameter in both the CD and the ID model. This is because in both models, the value of the outside option that enters the participation constraint depends on income persistence. In the ID model this outside option is autarky, which becomes more attractive for high income households when incomes are more persistent, implying that future incomes are expected to also be high. At the same time, higher persistence reduces the probability of receiving transfers in the future. Together, this makes a deviation more attractive for high income individuals when persistence is high. In the CD model there is an additional effect, as the persistence of incomes changes the relative attractiveness of high vs. low income villagers as partners in a deviating coalition. Thus, when incomes are persistent, those with high income today are attractive partners as they are likely to also be income-rich tomorrow, and thus able to provide consumption resources for others with low income draws. With negative serial correlation, the same is true for villagers with low income (who, however, may not be willing to deviate as they are receivers of transfers today, and are expected to be income rich only 'every other' period).

In our benchmark calibration, we estimated the income process directly from the data. This resulted in persistence that was small in absolute magnitude for all villages. We argued why we chose this benchmark process for incomes, based on an estimation that allowed households to differ in their mean incomes, but assumed common persistence. Alternative approaches, such as an estimation for different income groups (Laczó, 2014), household-specific income processes (Ligon et al., 2002), or a homogeneous income process for all villagers (as in Section A.14), have also been used, and imply different estimates of income persistence. In this section we study whether the standard ID model is able to capture the observed income-consumption distribution better when we choose the value of the persistence parameter ρ that best matches the observed consumption-income comovement in data. Intuitively, one might expect negative serial correlation in incomes to reduce the asymmetry of the joint distribution of consumption and income in that model by flattening the relationship between income and autarky values. For each of the three villages, and given a common village-specific cross-sectional dispersion of incomes that we take from the data, we therefore estimate the value of the persistence parameter ρ freely on a grid between -0.6 and $+0.6$ to best match our four target moments.

Table A.12 reports the results. The CD model predicts increased persistence relative to the estimates in Table 3, particularly in Aurepalle, where it also predicts a higher discount factor, which interact to leave the predicted moments approximately unchanged. The fact that serial correlation does not much affect the moments predicted by the CD model is also evidenced by the large uncertainty surrounding the estimate of ρ and δ. At an estimated value of ρ equal to the lower bound of -0.6 in all three villages, and a discount factor δ substantially lower than in our benchmark results in Table 5, the ID model now predicts a degree of insurance that is lower, and thus closer to the data, at the price of a somewhat increased asymmetry. The overall goodness of fit is only marginally improved, however. So even with a serial correlation that is counterfactually negative relative to the data, the ID model is not able to simultaneously predict a symmetric distribution of consumption and income growth and a realistic observed degree of insurance.

A.13. Preference heterogeneity in the ID model

The estimates in Table 5 suggested that the benchmark specification of the ID model, where all households were assumed to have identical preferences, was not able to explain the degree of risk sharing observed in the data at the same time as approximately symmetric consumption-income comovement. Heterogeneity in preferences, in contrast, is, rather trivially, able to reconcile a realistic degree of insurance with symmetry in consumption in the ID model. This is because both autarkic allocations (where households consume their income) and full insurance imply symmetry. The right mix of approximately risk-neutral and highly risk-averse households, experiencing perfect and zero income-consumption comovement respectively, can thus always deliver the right average comovement and symmetry in our moments-based approach. With a less extreme degree of preference heterogeneity, in contrast, the asymmetry predicted by the standard model may actually be stronger than in the benchmark. This is because when insurance is strong, as in the ICRISAT data, introducing dispersion in, for example, risk aversion around its estimated mean moves more risk-averse households even closer to perfect insurance, which may not change their consumption moments much. Insurance declines, however, for less risk-averse households, who care less about insurance and whose participation constraints thus bind more often, and whose consumption declines faster when they are

TABLE A.12. Preferences and persistence estimated to target degrees of risk sharing and asymmetry

	Aurepalle			Kanzara			Shirapur		
	Data	**CD**	**ID**	**Data**	**CD**	**ID**	**Data**	**CD**	**ID**
n		4.00	34.00		4.00	37.00		4.00	31.00
δ		0.98	0.74		0.98	0.84		0.98	0.83
s.e.		6.15	0.12		352.47	0.08		643.29	0.12
σ		1.00	1.00		1.00	1.00		1.00	1.00
$\frac{Var_{dc}}{Var_{dy}}$	0.30	0.24	0.08	0.56	0.23	0.08	0.33	0.23	0.03
β_{dcdy}	0.21	0.23	0.20	0.22	0.23	0.20	0.17	0.23	0.10
$\frac{Var_{dc\vert dy>0}-Var_{dc\vert dy\leq 0}}{Var_{dy}}$	-0.08	-0.00	0.10	-0.25	-0.00	0.10	-0.16	-0.00	0.04
$\beta_{dcdy>0}-\beta_{dcdy\leq 0}$	-0.41	-0.00	0.24	-0.14	-0.00	0.24	-0.05	-0.00	0.15
ρ		0.60	-0.60		0.20	-0.60		0.20	-0.60
s.e.		2.09	0.45		8.31	0.58		5.75	0.73
Model criterion		12.10	48.52		5.45	14.80		6.30	26.77

Note: For each of the three ICRISAT villages, the table shows four key moments summarising the joint distribution of consumption and income growth in the survey ("Data", in the first column for each village), and in simulations of the CD model (in the second column for each village) and the ID model (third column). For the simulated model solutions, apart from the size of the insurance groups n, the table also presents the discount factor δ and the value of income persistence, which are chosen to minimise the sum of differences between all four moments predicted by the models and those observed in the data, weighted by the inverse variance of the data moments calculated using a bootstrap procedure. The estimated parameters are those that minimise this criterion on a grid of $\delta \in [0.5, 0.98]$ and $\rho \in [-0.6, 0.6]$. The last line reports the value of the criterion at the estimated parameters.

unconstrained. This implies a decrease in insurance, and an increased asymmetry in their consumption process (as they move up the right hand side of the inverse U-shaped relation between asymmetry and the degree of risk sharing). At the strong insurance predicted by the standard model, this increase in the asymmetry for the less risk-averse may dominate the decline for more risk-averse households (and the increased dispersion of consumption growth for the unconstrained).

Table A.13 compares our benchmark estimates of the ID model to an alternative where villages consist of an equal number of villagers with risk aversion coefficients equal to 0.5 and 1.5 (implying a number of villagers divisible by two, which reduces those in Kanzara and Shirapur by 1). The fall in insurance implied by the presence of agents that are substantially less risk averse is counteracted by an increase in the estimated discount factor δ in all villages. The heterogeneity acts to increase the relative variance of consumption growth, while the increased discount factor decreases the regression coefficient β_{dcdy} and the corresponding asymmetry moments. Overall, the estimates are changed little by the inclusion of this stylised form of heterogeneity.

Table A.13. Risk sharing moments for heterogeneous preferences

	Aurepalle			Kanzara			Shirapur		
	Data	**ID**	**ID Het**	**Data**	**ID**	**ID Het**	**Data**	**ID**	**ID Het**
n		34.00	34.00		37.00	36.00		31.00	30.00
δ		0.89	0.92		0.90	0.92		0.88	0.92
s.e.		0.01	0.01		0.01	0.02		0.01	0.01
σ		1.00	1.50		1.00	1.50		1.00	1.50
$\frac{Var_{dc}}{Var_{dy}}$	0.30	0.04	0.06	0.56	0.07	0.09	0.33	0.02	0.03
β_{dcdy}	0.21	0.11	0.11	0.22	0.17	0.16	0.17	0.07	0.07
$\frac{Var_{dc\|dy>0}-Var_{dc\|dy\leq 0}}{Var_{dy}}$	-0.08	0.07	0.09	-0.25	0.10	0.12	-0.16	0.03	0.04
$\beta_{dcdy>0}-\beta_{dcdy\leq 0}$	-0.41	0.20	0.19	-0.14	0.28	0.25	-0.05	0.13	0.13
Model criterion		52.76	50.42		15.89	15.11		27.17	26.15

	Aurepalle			Kanzara			Shirapur		
	Data	**ID**	**ID Het**	**Data**	**ID**	**ID Het**	**Data**	**ID**	**ID Het**
n		34.00	34.00		37.00	36.00		31.00	30.00
δ		0.89	0.92		0.90	0.92		0.88	0.92
s.e.		0.01	0.01		0.01	0.02		0.01	0.01
σ		1.00	1.50		1.00	1.50		1.00	1.50
$\frac{Var_{dc}}{Var_{dy}}$	0.30	0.04	0.06	0.56	0.07	0.09	0.33	0.02	0.03
β_{dcdy}	0.21	0.11	0.11	0.22	0.17	0.16	0.17	0.07	0.07
$\frac{Var_{dc\|dy>0}-Var_{dc\|dy\leq 0}}{Var_{dy}}$	-0.08	0.07	0.09	-0.25	0.10	0.12	-0.16	0.03	0.04
$\beta_{dcdy>0}-\beta_{dcdy\leq 0}$	-0.41	0.20	0.19	-0.14	0.28	0.25	-0.05	0.13	0.13
Model criterion		52.76	50.42		15.89	15.11		27.17	26.15

Note: For each of the three ICRISAT villages, the table shows the four key moments summarising the joint distribution of consumption and income growth in the survey ("Data", in the first column for each village), and in simulations of the ID model (in the second column for each village) and a simple form of preference heterogeneity, where the village population comprises two groups of equal size whose risk aversion σ equals 0.5 and 1.5 respectively.

Table A.14. Estimated income process without household fixed effects

	Aurepalle	**Kanzara**	**Shirapur**
ρ	0.72	0.76	0.59
Var_ε	0.28	0.15	0.25

Note: The table presents the point estimates for the persistence parameter ρ and the shock variance Var_ε for the AR(1) process (27) without household fixed effects.

A.14. Estimating the model and income process without fixed effects

Table A.14 reports estimates of the income process (27) without household fixed effects. Table A.15 reports the model estimates for this case, taking all four moments as targets. The income process is substantially more persistent and more volatile (as it mechanically attributes any permanent intra-household variation to persistent shocks), particularly for Kanzara and Shirapur. The increase in persistence and volatility have offsetting effects on the value of autarky at high income (as more persistent high income increases the attractiveness of autarky, while more volatile income shocks reduce it). Together with an increase in the estimated discount factor, this leaves the results for both the ID and CD models little changed (although the estimated group size in Kanzara is reduced by one household). Particularly, at the higher discount factor, the CD model continues to predict close-to-full insurance. The overall conclusion is therefore the same: no model can replicate the negative asymmetry in Aurepalle, but the CD model performs substantially better in replicating both the degree and approximate symmetry in all other villages. In fact, the overall fit of the CD model is improved compared to the benchmark estimates in Table 5, both in absolute terms and relative to the other comparison models.

A.15. Measurement error in consumption and incomes

We now examine the extent to which the introduction of measurement error may change our results. Measurement error in consumption may help the models to reconcile a sizeable volatility of measured consumption with a modest comovement of consumption and income. Measurement error in incomes, on the other hand, has three main effects on the moments we focus on: first, it attenuates the slope coefficient in a regression of consumption growth on income growth. Second, when more of the measured variance of incomes comes from random error, the true income process becomes less variable and its autocorrelation increases in absolute magnitude. This is because in order to predict a given measured autocovariance, which is unaffected by classical measurement error, less volatile income shocks have to be more persistent. Finally, measurement error "blurs" the asymmetry moments that condition on the sign of income changes. Whenever the latter is mainly determined by measurement error, we would thus expect consumption moments to be identical for households with rising and falling incomes.

In this section we generalise our model by assuming that measured consumption, as well as income, include measurement error given by

$$\widehat{c_{it}} = c_{it} + \xi_{it} \tag{A.2}$$

$$\widehat{y_{it}} = y_{it} + \nu_{it} \tag{A.3}$$

where $\widehat{x_{it}}$ and x_{it} are the measured and true levels of the logarithm of variable x, and ξ_{it} and ν_{it} denote measurement error in consumption and income that is identically and independently distributed across individuals and time with variance Var_{ξ} and Var_{ν}.

Note that, for a given measured variance and autocovariance of income in the data $Var_{\widehat{y}}$ and $Cov_{\widehat{y}}$, the persistence parameter ρ and variance of 'true' income shocks Var_{ε}, which are both an input to the model, are now a function of Var_{ν}:

$$\rho = \frac{Cov_{\widehat{y}}}{Var_{\widehat{y}} - Var_{\nu}} \tag{A.4}$$

$$Var_{\varepsilon} = (Var_{\widehat{y}} - Var_{\nu}) * (1 - \rho^2). \tag{A.5}$$

This necessity of solving the model afresh for all parameters, villages, and both models when Var_{ν} takes a new value constrains us to a small number of values for Var_{ν}. We therefore constrain measurement error by allowing the variance of measured income growth to be at most three times that of true income y_{it}. Also note that, for a given measured variance and autocovariance of incomes $Var_{\widehat{y}}$ and $Cov_{\widehat{y}}$, an increase in measurement error increases the persistence of income shocks ρ, as mentioned above.

Table A.16 presents the results. Estimates of measurement error in both consumption and income are substantial in all models. Specifically, measurement error in consumption, which accounts for between 30 and 94 percent of the variance of measured consumption growth, allows all models to fit the relative variance of consumption and income growth Var_{dc}/Var_{dy} almost perfectly.[33] The improved fit comes at the cost of reducing the empirical content of the models: since consumption measurement error leaves all other moments unchanged, it effectively removes the relative variance of consumption and income growth from the objective function.

Measurement error in incomes is also estimated to be substantial and equals the maximum value of our grid, where it accounts for two thirds of the variance in measured income growth, in 7 of the 9 cases we consider. This results in a strongly improved fit of the comparison models. In the SI model, in particular, it reduces the asymmetry and attenuates the regression coefficient β_{dcdy}, bringing it in line with values observed in the data. Measurement error also reduces the asymmetry in the ID model, which, however, continues to predict a counterfactually low coefficient β_{dcdy}.

Generally, the estimates in Table A.16 imply that the joint distribution of consumption and income is dominated by measurement error, particularly in the case of the ID model. Apart from a counterfactually low coefficient β_{dcdy} this implies a counterfactually low fit of the consumption data in a Townsend (1994)-type regression of individual consumption growth on individual income growth and time fixed effects. Specifically, the R^2 of this regression in the ICRISAT data is 0.40, 0.31 and 0.20 for, respectively, Aurepalle, Kanzara and Shirapur. In contrast, that predicted by the ID model is 0.063, 0.067 and 0.081.

There is an additional reason why the results in Table A.16, particularly for the CD and ID models, should be treated with caution, namely the extreme uncertainty surrounding several of the point estimates. In the ID model, for example, the discount factor δ is not identified in the case of Kanzara (where the model predicts autarky plus strong attenuation, and predicted moments are thus unchanged by small changes in δ). In the CD model, while consumption measurement error is clearly identified by the relative volatility of consumption and incomes, income measurement error is often not identified, as illustrated by the extreme standard errors surrounding some of its estimates. This is

33. We constrain consumption measurement error to lie on a grid that takes 31 values, hence the small difference between the data and model moments.

because the relative volatility of consumption and incomes is fitted perfectly by measurement error in consumption and the asymmetry moments are small and only little affected by income measurement error and changes in the discount factor. The regression coefficient β_{dcdy} can then often be matched in two ways: with a high discount factor, implying strong insurance in larger groups of size 4 and 5, and zero measurement error (as seen in Section 4 and here in Table A.16 for Shirapur). Or with substantial measurement error, which reduces the true volatility of incomes, making membership in large insurance groups less attractive, and thus raising the 'fundamental' regression coefficient by reducing group size (as seen here for Kanzara). A further increase in measurement error can then often attenuate the measured β_{dcdy} to match the data. In fact, in our solution maximum group sizes decline monotonically with income measurement error, to 2 or 3 households at the high values estimated in Table A.16. Our approach, which uses a grid of measurement error and discount factors does not capture this lack of identification exactly. But the goodness of fit in Table A.16 is very similar to that of the CD model with measurement error only in consumption, equal to 10.8, 2.1 and 4.0, with maximum stable group sizes of 4, 4, and 6, in Aurepalle, Kanzara and Shirapur respectively. These group sizes can be viewed as an upper bound on the largest sustainable sizes, which would be reduced by measurement error in incomes, for which the four moments we focus on are, however, a poor guide in the case of the CD model.[34]

34. It is tempting to compare these numbers to those in Table A.16 using a Chi-Sqare distribution. But since the estimated income measurement error equals the bound of our grid, this would be invalid.

TABLE A.15. Preferences estimated to target all 4 moments (no fixed effects)

	Aurepalle				Kanzara				Shirapur			
	Data	**CD**	**ID**	**SI**	**Data**	**CD**	**ID**	**SI**	**Data**	**CD**	**ID**	**SI**
n		4.00	34.00			3.00	37.00			4.00	31.00	
δ		0.98	0.91	0.94		0.98	0.93	0.95		0.98	0.90	0.95
s.e.		0.00	0.01	0.01		0.01	0.01	0.00		0.00	0.01	0.01
σ		1.00	1.00	1.00		1.00	1.00	1.00		1.00	1.00	1.00
$\frac{Var_{dc}}{Var_{dy}}$	0.30	0.26	0.03	0.11	0.56	0.35	0.07	0.12	0.33	0.24	0.04	0.07
β_{dcdy}	0.21	0.22	0.08	0.27	0.22	0.31	0.13	0.29	0.17	0.22	0.09	0.20
$\frac{Var_{dc\|dy>0}-Var_{dc\|dy\leq 0}}{Var_{dy}}$	-0.08	0.01	0.07	0.09	-0.25	0.00	0.14	0.09	-0.16	0.00	0.06	0.05
$\beta_{dcdy>0}-\beta_{dcdy\leq 0}$	-0.41	-0.00	0.15	0.15	-0.14	0.00	0.24	0.14	-0.05	0.00	0.16	0.12
Model criterion		11.46	52.86	37.97		4.75	16.36	11.38		5.63	27.66	20.56

	Aurepalle				Kanzara				Shirapur			
	Data	**CD**	**ID**	**SI**	**Data**	**CD**	**ID**	**SI**	**Data**	**CD**	**ID**	**SI**
n		4.00	34.00			3.00	37.00			4.00	31.00	
δ		0.98	0.91	0.94		0.98	0.93	0.95		0.98	0.90	0.95
s.e.		0.00	0.01	0.01		0.01	0.01	0.00		0.00	0.01	0.01
σ		1.00	1.00	1.00		1.00	1.00	1.00		1.00	1.00	1.00
$\frac{Var_{dc}}{Var_{dy}}$	0.30	0.26	0.03	0.11	0.56	0.35	0.07	0.12	0.33	0.24	0.04	0.07
β_{dcdy}	0.21	0.22	0.08	0.27	0.22	0.31	0.13	0.29	0.17	0.22	0.09	0.20
$\frac{Var_{dc\|dy>0}-Var_{dc\|dy\leq 0}}{Var_{dy}}$	-0.08	0.01	0.07	0.09	-0.25	0.00	0.14	0.09	-0.16	0.00	0.06	0.05
$\beta_{dcdy>0}-\beta_{dcdy\leq 0}$	-0.41	-0.00	0.15	0.15	-0.14	0.00	0.24	0.14	-0.05	0.00	0.16	0.12
Goodness of fit		11.46	52.86	37.97		4.75	16.36	11.38		5.63	27.66	20.56

Note: For each of the three ICRISAT villages, the table shows four key moments summarising the joint distribution of consumption and income growth in the raw data of the survey ("Data", in the first column for each village), and in simulations using the income process estimated with unconditional data (see Table A.10) in the CD model (in the second column for each village), the ID model (third column) and the SI model (fourth column). For the simulated model solutions, the table also presents the size of the insurance groups n (in the case of the ID and CD models) and the discount factor δ, which is chosen to minimise the sum of differences between all four moments predicted by the models and those observed in the data, weighted by the inverse variance of the data moments calculated using a bootstrap procedure. The estimated preference parameters are those that minimise this criterion on a grid of $\delta \in [0.5, 0.99]$. The last line reports the value of the criterion at the estimated parameters.

Table A.16. Preferences estimated to target all 4 moments and measurement error

	Aurepalle				Kanzara				Shirapur			
	Data	**CD**	**ID**	**SI**	**Data**	**CD**	**ID**	**SI**	**Data**	**CD**	**ID**	**SI**
n		4.00	34.00			2.00	37.00			5.00	31.00	
δ		0.97	0.95	0.91		0.96	0.50	0.88		0.97	0.85	0.82
s.e.		39.80	0.37	0.07		16.71	486.76	0.19		354.59	1.08	0.18
σ		1.00	1.00	1.00		1.00	1.00	1.00		1.00	1.00	1.00
φ												
$\frac{Var_{dc}}{Var_{dy}}$	0.30	0.30	0.30	0.30	0.56	0.57	0.56	0.56	0.33	0.33	0.35	0.33
β_{dcdy}	0.21	0.20	0.02	0.22	0.22	0.22	0.15	0.23	0.17	0.17	0.09	0.17
$\frac{Var_{dc\|dy>0}-Var_{dc\|dy\leq 0}}{Var_{dy}}$	-0.08	-0.00	0.01	0.04	-0.25	-0.00	0.09	0.03	-0.16	-0.00	0.03	0.01
$\beta_{dcdy>0}-\beta_{dcdy\leq 0}$	-0.41	-0.01	0.04	0.06	-0.14	-0.00	0.11	0.05	-0.05	-0.00	0.03	0.02
$\frac{Var_{d\varepsilon}}{Var_{dc}}$		0.30	0.94	0.45		0.61	0.80	0.69		0.44	0.85	0.72
s.e.		0.32	0.64	0.35		12.16	0.08	0.07		1.77	0.09	0.24
$\frac{Var_{d\nu}}{Var_{dy}}$		0.12	0.64	0.64		0.55	0.59	0.65		0.00	0.65	0.65
s.e.			7.48	0.38		19.45	0.64	0.51			0.19	0.15
Goodness of fit		10.25	21.42	15.46		1.70	5.12	2.53		3.55	6.89	4.57

Note: For each of the three ICRISAT villages, the table shows four key moments summarising the joint distribution of consumption and income growth in the survey ("Data", in the first column for each village), and in simulations of the CD model (in the second column for each village), the ID model (third column) and the SI model (fourth column). For the simulated model solutions, the table also presents the size of the insurance groups n and the discount factor δ, and measurement error in consumption and income $Var_{d\varepsilon}/Var_{dc}$ and $Var_{d\nu}/Var_{dy}$. δ, $Var_{d\varepsilon}/Var_{dc}$ and $Var_{d\nu}/Var_{dy}$ are chosen to minimise the sum of differences between all four moments predicted by the models and those observed in the data, weighted by the inverse variance of the data moments calculated using a bootstrap procedure. The estimated parameters are those that minimise this criterion on a grid of $\delta \in [0.5, 0.99]$. The last line reports the value of the criterion at the estimated parameters. The discount factors associated with the ID model in Kanzara are not identified (N.I.), as the model predicts autarky there.