Risk Pooling and Precautionary Saving in Village Economies

Marcel Fafchamps and Aditya Shrinivas

May 2022

Abstract

We propose a new method to test for efficient risk pooling that allows for intertemporal smoothing, non-homothetic consumption, and heterogeneous risk and time preferences. The method is composed of three steps. The first one allows for precautionary savings by the aggregate risk pooling group. The second utilizes the inverse Engel curve to estimate good-specific tests for efficient risk pooling. In the third step, we obtain consistent estimates of households' risk and time preferences using a full risk sharing model, and incorporate heterogeneous preferences in testing for risk pooling. We apply this method to panel data from Indian villages to generate a number of new insights. We find that food expenditures are better protected from aggregate shocks than non-food consumption, after accounting for non-homotheticity. Village-level consumption tracks aggregate village cash-in-hand, suggesting some form of coordinated precautionary savings. But there is considerable excess sensitivity to aggregate income, indicating a lack of full asset integration. We also find a large unexplained gap between the variation in measured consumption expenditures and cash-in-hand at the aggregate village level. Contrary to earlier findings, risk pooling in Indian villages no longer appears to take place more at the sub-caste level than at the village level.

1 Introduction

People in village economies face a large number of income shocks due to drought, flood, unemployment, illness, and crop or business failure. Households that are uninsured against these shocks experience consumption fluctuations detrimental to their welfare (Gertler & Gruber, 2002; De Weerdt & Dercon, 2006). Protection from such income shocks depends on the availability and effectiveness of institutions that distribute and share risk. Asset accumulation *ex ante* can help smooth consumption through precautionary saving (e.g., Zeldes (1989), Deaton (1991)). Risk can also be pooled *expost* through various informal or formal agreements (e.g., Udry (1990), Fafchamps & Lund (2003), Dercon et al. (2006)). Risk pooling is best at addressing idiosyncratic shocks that affect only some households at a time. But it offers little or no protection against aggregate shocks that affect the whole village. To self-protect against such shocks, some form of precautionary saving is required – either at the aggregate level (e.g., cereal bank; provident fund) or at the individual level (Fafchamps et al., 1998).

In this paper we examine detailed panel data from thirty villages in India for evidence of risk pooling and precautionary saving. We propose three methodological extensions to the standard test of risk pooling. First, we allow for precautionary savings and the accumulation of assets within that model. To the best of our knowledge, this is the first study that combines both aspects of risk pooling and precautionary savings in an integrated and theoretically-based framework. Second, we expand the standard model to allow testing for risk pooling for different categories of consumption expenditures while simultaneously allowing for non-homothetic consumption preferences. Third, in accordance to recent advances in testing for risk pooling, we correct for differences in risk and time preferences among households. We implement an original methodology for estimating risk pooling tests that account for heterogeneity in risk and time preferences. This method is based on the intuition, dating back to Wilson (1968), that efficient risk pooling allocates more risk to more risk-tolerant households, implying that a household whose consumption strongly co-moves with village consumption must be relatively more risk tolerant. Finally, we replicate the analysis assuming aggregation of risk either at the village-level or the sub-caste level.

For total expenditures, we find evidence of less than perfect risk pooling at the village level. But the shortfall is quite limited in the sense that household consumption expenditures move nearly one-for-one with the village average. Even when year-to-year variation in household-level liquid wealth and earned income have a statistically significant coefficient, the magnitude of the correlation with household expenditures remains very small. From this we conclude that the study villages engage in a considerable pooling of idiosyncratic within-year risk. We are not saying, however, that risk pooling is the result of the explicit sharing of risk among villagers. Indeed, risk pooling tests do not identify the mechanism by which risk is being pooled.[1] This reality, which is shared by all risk pooling test papers in the literature, is the reason why, throughout the paper, we have refrained from using the expression 'risk sharing' since it implicitly suggests some deliberate sharing intent.

We find that food and non-food expenditures do not vary with average village consumption in a way consistent with a within-year utility-maximizing allocation of total expenditures: households increase their food consumption less – and their non-food consumption more – than what the increase in average village expenditures would predict, based on estimated non-linear Euler curves. We also do not find that non-food expenditures respond more to household income and wealth than non-food expenditures, ruling out the idea that there is less risk pooling in non-food consumption than in food consumption. Instead, the results suggest that, in a good year for the village as a whole, people increase their non-food expenditures more than proportionally, thereby sheltering their average food consumption from aggregate shocks.

Our results confirm that there is substantial and significant heterogeneity in estimated risk and time preferences across households. We also find, as previous authors have argued theoretically, that the failure to correct for this heterogeneity biases coefficient estimates in any risk pooling test. In our data, however, this bias is empirically negligible and it does not affect the qualitative conclusions from the analysis.

Turning to consumption smoothing across years, we find a small but significant response of village consumption to aggregate village cash-in-hand. This finding is consistent with a high level of consumption smoothing achieved through precautionary saving – i.e., close to certainty equivalence – but it is partially negated by the larger coefficient on village earned income, which indicates a failure of asset integration: villages do not appear to optimally draw on liquid assets to smooth aggregate income shocks. Results also indicate that village food consumption, suitably corrected for non-homothetic preferences, responds little to cash-in-hand while village non-food expenditures respond significantly to cash-in-hand and display excess income sensitivity.

Similar findings are obtained when the same analysis is replicated at the sub-caste (*Jati*) level, instead of the village level. But they also suggest that, contrary to the existing literature (e.g., Townsend (1994) (Mazzocco & Saini, 2012; Shrinivas & Fafchamps, 2018)), risk pooling within sub-castes with each village is less strong – i.e., less responsive to aggregate consumption and more responsive to individual income and liquid assets – than pooling across all households in the village. Conclusions are also similar regard-

[1] Risk pooling can of course be achieved through an explicit or implicit agreement to share risk. However, as already noted in Sargent 1990 (chapter 3), a large amount of risk pooling can be achieved through individual precautionary saving. Sargent demonstrates this with the simple example of an island economy in which individual producers use their stock of currency to buy food when in deficit and sell food when in surplus. As long as they do not run out of currency, producers can achieve near perfect smoothing of consumption by using precautionary saving to pool risk. Only when they stock out of liquid wealth do they experience an individual fall in consumption (e.g., Deaton (1992)

ing precautionary saving: aggregate sensitivity to cash-in-hand is very small – probably smaller than it should be given the relatively low level of liquid wealth in the data: in a poor population such as the one we study, we would expect a stronger dependence of consumption to cash-in-hand in a precautionary savings optimum. The sub-caste evidence also suggests that the excess sensitivity to income found in the linear (CARA) model may be due to misspecification; a log model provides a better fit for the theory.

This paper makes several contributions to the literature. First, our empirical analysis of risk pooling in village economies is the first to combine intra-temporal and inter-temporal aspects of smoothing in a theoretically consistent framework. Previous studies of consumption smoothing across individuals have focused on risk pooling within periods (e.g., Mace (1991), Altonji et al. (1992), Townsend (1994), (Schulhofer-Wohl, 2011; Mazzocco & Saini, 2012; Chiappori et al., 2014) – and those that have examined risk pooling across periods have ignored or assumed away assets (e.g., Kocherlakota (1996), Ligon (1998), Ligon et al. (2002),Kinnan (2021)). On the other hand, past studies of intertemporal risk coping via precautionary savings have typically assumed away contemporaneous risk pooling across households (e.g., Zeldes (1989), Deaton (1991), Rosenzweig & Wolpin (1993), Lim & Townsend (1998), Fafchamps et al. (1998),Kazianga & Udry (2006)).

We show how the two empirical approaches can be combined within a single framework. Past studies of risk pooling are nested in our model and, as such, are unbiased. In contrast, studies of precautionary saving that ignore risk pooling across individuals may yield misleading results. This is because, in the presence of risk pooling, individually held assets can be used to smooth the consumption of others. In an efficient equilibrium, all assets are, de facto, held in trust for the entire risk pooling group. It follows that the efficient distribution of precautionary assets across households only depends on whether returns are convex, concave, or linear: if they are convex, all precautionary assets should be held by a single actor (e.g., an insurance corporation); if they are concave, they should be equally distributed; and if they are linear, the distribution of assets does not matter. Hence precautionary saving can only be meaningfully studied at the level of the risk pooling group aggregate.

Second, this paper contributes to the new strand of literature on risk sharing with heterogeneous preferences (Schulhofer-Wohl, 2011; Mazzocco & Saini, 2012; Chiappori et al., 2014; Dubois, 2000). This literature has brought to light the existence of an omitted variable bias in standard tests of risk pooling under homogeneous preferences, and has shown that this bias drives the income coefficient upwards, leading to spurious rejections of full insurance. Different parametric and semi-parametric tests have been proposed to account for heterogeneous risk preferences. We follow in the footsteps of this literature by improving on the testing approach proposed by Chiappori et al. (2014). We start by estimating relative risk preferences between households under the assumption of perfect risk sharing and then use the resulting estimates to correct the test of full risk pooling. Our approach, however, differs from Chiappori et al. (2014) in important ways: it generalizes the approach by simultaneously estimating relative time preferences between households under perfect risk pooling; and it is easier to implement since in relies on a simple linear regression.

Third, we integrate non-homothetic Engel curves into our testing strategy. There has been a recent revival of interest in Engel curves (Dunbar et al., 2013; Atkin et al., 2020; Ligon, 2020; Almås et al., 2018; Escanciano et al., 2021). We propose a new way of using partial consumption data to test for efficient risk pooling when income elasticities are non-unitary. This approach also allows us to test whether certain components of consumption are better insured than others (e.g., Mace (1991)) – as could arise, for instance, in the presence of paternalism or imperfect altruism.[2]

Lastly, we make an empirical contribution to the literature on efficiency in risk sharing. Earlier results have emphasized that full risk sharing is not achieved in village economies. Using the new wave of ICRISAT's panel data from 2010-2015, our results are qualitatively consistent with previous estimates from the 1975-85 ICRISAT data: household income is statistically significant. We nonetheless find that the marginal propensity to consume (MPC) out of household's own income is a very low 0.02, which is substantially lower than the MPC of 0.14 estimated by Townsend (1994) using the old data. In absolute terms, Rosenzweig & Binswanger (1993) find that a 100 rupee decline in profits reduces food consumption by 7 rupees in the old sample. In the new sample, an equivalent decline in earned income reduces food consumption by a mere 3.4 rupees. These results suggest that consumption smoothing has improved in these villages over time. In addition, our findings corroborate recent papers that find substantial heterogeneity in MPC across households (Lewis et al., 2019; Aguiar et al., 2020). Earlier work based on the 1976-85 data indicated a higher level of risk pooling within sub-castes. Our findings from the more recent data show little difference between village and sub-caste results: sensitivity to cash-in-hand is very small in both cases – and possibly smaller within sub-castes.

We however go well beyond existing studies in the scope of our findings. While considerable risk pooling occurs within villages or sub-castes, sheltering households against idiosyncratic shocks, our evidence suggests that villages achieve very little protection across aggregate year-to-year shocks via precautionary savings: group consumption hardly responds at all to liquid village-level liquid assets or cash-in-hand. We nonetheless find that aggregate village consumption also responds little to aggregate village income. Taken together, these findings suggest that average village consumption is largely insulated from village-specific income shocks – although the process by which this is achieved is unclear, apart from the fact that it is not due to precautionary savings. One possibility, discussed

[2]The long-lasting US Food Stamp program is a well-known example where one dimension of consumption – i.e., food – is better insured than others – e.g., utilities, heating, transport (e.g., Hastings & Shapiro (2018)).

in the literature is transfers within sub-caste across villages (e.g., Rosenzweig & Stark (1989)); another is transfers from migrant workers (e.g., Munshi & Rosenzweig (2016)).

While our empirical analyis largely abstracts away from studying the mechanisms that limit full risk pooling,[3] we explore one little-explored source of friction, namely, the deviation of consumption expenditures from what would be optimal for the household. This could arise, for instance, because consumption categories are not all equally observable. As a result, individuals may claim a negative shock but spend the insurance payout on luxuries. To deter such behavior, recipients of an insurance payouts may be forced to spend it on necessities such as food. We do indeed find evidence that food consumption is over-insured in the sense that it co-moves less with village consumption than is predicted by Engel curves, while non-food consumption co-moves more. This suggests that if households were free to spend their consumption budget freely, they would opt to spend less on food and more on non-food in bad years.

2 Risk pooling with Assets: A conceptual framework

Since we are interested in testing risk pooling within villages, we follow Townsend (1994), and many others and assume a closed village economy over time. In each period t each individual i in the village receives an earned income y_{its} that varies with the state of nature s. We assume that the joint income distribution of all individuals in the village is stationary over time with known mean, variance, and covariance vectors. This allows correlation in outcomes across individuals within period but, for simplicity, we abstract from autocorrelation of incomes across time.[4] The probability of state of the world s is denoted π_s. Each individual starts the period with liquid wealth $(1 + r_s)w_{it}$ where r_s is the return to assets, which is allowed to vary with the state of the world s. Each

[3] Several explanations have been proposed for the failure of full risk pooling. One is moral hazard: when risk taking behavior is not observable to others, being insured may lead to excessive risk taking; this in turn may induce caps on insurance coverage (Rogerson, 1985; Golosov et al., 2003). Another is limited commitment: households receiving a high income draw may leave the insurance arrangement when asked to make a large contribution to the insurance pool; this limits the contributions households can be expected to make (Kimball, 1988; Coate & Ravallion, 1993; Ligon et al., 2002; Laczó, 2015). A third possibility is hidden income, opening the possibility of insurance fraud; this in turn results in some forms of income risk being better insured than others (Townsend, 1982). Kinnan (2021) examines several of these possibilities in a joint nested model: using panel data from rural Thailand, the author provides evidence that supports the hidden income hypothesis but rejects limited commitment and moral hazard. Furthermore, there is also an extensive literature on the role of network structures on risk-sharing (Ambrus et al., 2014; Ambrus & Elliott, 2020; Ambrus et al., 2021). For instance, (Ambrus et al., 2014) show that the degree of risk-sharing is governed by the expansiveness of the network, and households consumption will co-move more strongly with that of socially closer households. (Ambrus et al., 2021) study local information constraints in risk-sharing networks and predict that network centrality is positively correlated with consumption volatility. Similarly, (Chandrasekhar et al., 2018) show that the well-connected central agents engage more in risk-sharing networks when income risk is high, when income shocks are positively correlated and when attitudes towards risk are more sensitive in the aggregate.

[4] Differences in the mean of income across individuals get subsumed in the welfare weights.

individual's cash-in-hand at the beginning of the period is thus $x_{its} \equiv y_{its} + (1 + r_s)w_{it}$. We restrict our attention to cases where the total liquid wealth of the village must be non-negative. But the individual net liquid wealth of individuals can be negative.

The utility that an individual derives from consumption expenditures c_{its} is given by a standard instantaneous utility function $U_i(c_{its})$ specific to individual i. This allows for heterogeneous risk preferences. Each individual discounts the future with constant discount factor ρ_i, which similarly allows for heterogeneous time preferences.

We identify the Pareto efficient allocation of consumption across individuals within and across periods by solving a social planner problem of the form:

$$\begin{aligned}
&\underset{\{c_{its}, w_{its}\}}{\text{Max}} && \sum_{t=1}^{T}\sum_{i=1}^{N}\eta_i\rho_i^t\sum_{s=1}^{S}U_i(c_{its})\pi_s \\
&\text{s.t.} && \sum_{i}^{N}c_{its} = \sum_{i}^{N}\left((1+r_s)w_{it} + y_{its} - w_{it+1,s}\right) \quad \forall t, s \qquad (1)\\
& && \sum_{i}^{N}w_{it+1,s} \geq 0 \quad \forall t, s
\end{aligned}$$

where η_i is a particular set of (time-invariant) welfare weights with $\sum_{i=1}^{N}\eta_i = 1$. The middle equation (1) denotes the aggregate feasibility constraint that must hold in each time period t and state of the world s. To each particular set of welfare weights $\{\eta_i\}$ corresponds a different Pareto efficient solution.

We now characterize the solution to the social planner's problem. We begin by noting that any income vector that has the same aggregate income $y_{ts} = \sum_{i}^{N} y_{its}$ produces the same optimal solution. The same can be said for w_{ist}: any distribution of assets across individuals that generates the same total wealth $w_{ts} \equiv \sum_{i}^{N} w_{its}$ generates the same optimal solution. It follows that the allocation of consumption across individuals does not depend on individual incomes and wealth: only village aggregates y_{ts} and w_{ts} matter. Put differently, the social planner's problem satisfies income and asset pooling: within each period, individual welfare does not depend on individual income or assets realizations; rather, it depends on welfare weights and individual preferences. Second, we note that since the return to wealth is linear and identical across individuals, the way assets are distributed across individuals is irrelevant and thus undetermined. This means that the soluiton to the social planner's problem does not stipulate the distribution of liquid assets across individuals – only its aggregate.

Next we note that the social planner's problem can be decoupled into an inner optimization problem – how to allocate consumption across individuals, conditional on a choice of future savings $w_{t+1,s}$ for each s – and an outer optimization problem – how to allocate total consumption across periods by choosing the contingent path of $\{w_{t+1,s}\}$.

The inner optimization problem takes the familiar risk sharing form:

$$\begin{array}{lll} \underset{\{c_{its}|w_{t+1,s}\}}{\text{Max}} & \sum_{i=1}^{N} \eta_i \rho_i^t \sum_{s=1}^{S} U_i(c_{its}) \pi_s & \\ \text{s.t.} & \sum_{i}^{N} c_{its} = (1+r_s)w_t + y_{ts} - w_{t+1,s} \equiv c_{ts} \;\; \forall s & \end{array} \tag{2}$$

where $w_{t+1,s}$ is taken as given. Since $(1+r_s)w_t+y_{ts}$ is predetermined by past savings and the state of the world s, and $w_{t+1,s}$ is taken as given for the purpose of this optimization, the above optimization boils down to an allocation problem: how a given c_{ts} is divided among individuals. To characterize the properties of the solution, let us denote $\lambda_{ts}\pi_s$ as the Lagrange multiplier associated with the feasibility constraint. The first order conditions for the consumption levels c_{its} and c_{jts} of two arbitrary individuals in the same village are:

$$\eta_i \rho_i^t U_i'(c_{its}) = \lambda_{ts} = \eta_j \rho_j^t U_j'(c_{jts}) \tag{3}$$

which implies the usual condition for optimal risk pooling: since all individuals face the same realization of the aggregate resource constraint λ_{ts}, weighted marginal utilities of consumption are equated across individuals in each state of the world s. Since λ_{ts} is a deterministic function of c_{ts}, this leads to the standard testable prediction: individual consumption c_{its} varies with aggregate village consumption c_{ts}, not with individual income y_{ist} or wealth w_{ist}. This theoretical result forms the basis for all tests of efficient risk pooling.

We now turn to the outer optimization problem that selects the contingent aggregate level of savings $w_{t+1,s}$. Let $W_t(c_{ts})$ denote the value, to the social planner, of the *optimal* solution to the inner optimization problem for a total consumption level c_{ts}. Function $W_t(.)$ is indexed with t because, as we just discussed, when time preferences vary across individuals, the way the social planner divides the same amount of aggregate consumption c_{ts} across individuals varies over time. For clarity of exposition, let us define $R(t) \equiv \sum_{i=1}^{N} \eta_i \rho_i^t$. Since $\sum_{i=1}^{N} \eta_i = 1$ by construction, $R(t)$ is nothing but an average of individual discount factors weithed by the welfare weights.[5] Further, let us normalize individual discount factors as $\hat{\rho_i^t} = \frac{\rho_i^t}{R(t)}$ such that $\sum_{i=1}^{N} \eta_i \hat{\rho}_i^t = 1$. With this normalization, the outer optimization can be written in the form of the following Belman equation:

$$V_t(x_{ts}) = max_{w_{t+1,s}} W_t(x_{ts} - w_{t+1,s}) + R(t) E V_{t+1}((1+r_{s'})w_{t+1,s'} + y_{ts'})$$

where s' denotes the (yet unrealized) state of nature in period $t+1$ and where we made use of the fact that $c_{ts} = x_{ts} - w_{t+1,s}$. This is a standard optimization problem (Stokey

[5] Note that, as $t \to \infty$, $R(t)$ converges to the largest discount factor in the village.

and Lucas 1981). It yields as solution a policy function of the form $w_{t+1,s} = S_t(x_{ts})$. The case with homogenous time preferences has been extensively studied in the precautionary savings literature (e.g., Zeldes 1989, Deaton 1991). It is well know that $c_{ts} = C_t(x_{ts})$ is a concave function of x_{ts}.

2.1 Accounting for heterogeneous time and risk preferences

The generalization to heterogeneous time preferences does not change this main prediction. But the shape of $C_t(x_{ts})$ changes over time. This is because the relative weights associated with ratios of marginal utility vary over time: if i is more patient than j, then ρ_i^t/ρ_j^t increases with t. This means that i's expected share of aggregate consumption increases over time. This implies that early on, the social planer's discount factor $R(t)$ puts more weight on impatient individuals. As time passes, their weight in the average $\sum_{i=1}^{N} \eta_i \rho_i^t$ falls and $R(t)$ gets dominated by the most patient individuals whose weight ρ_i^t falls less fast. This means that as time passes, the marginal propensity to consume $\frac{\partial C(x_{is})}{\partial x_{is}}$ out of village assets falls. With infinitely lived agents, in the long run all village cash-in-hand x_{is} is consumed by the most patient individual(s) since the welfare weight $\eta_i \hat{\rho}_i^t$ of all the others converge more rapidly to 0. These are stark, unrealistic predictions that we do not expect to observe in practice, but they serve to outline the gradual unequalizing role that heterogeneous time preferences play in a risk pooling social optimum with assets.

Next we turn to the behavior of the model when individuals differ in their risk preferences. To this effect, we parameterize the utility function to have the constant-absolute-risk-aversion (CARA) form $U_i(c) = -\frac{e^{-\gamma_i c}}{\gamma_i}$ where parameter γ_i is the coefficient of absolute risk aversion of individual i.[6] With this functional form, the first order condition (3) simplifies to:

$$\eta_i \rho_i^t e^{-\gamma_i c_{its}} = \lambda_{ts}$$

Taking logs and rearranging yields, we get:

$$c_{its} = \frac{log\, \eta_i}{\gamma_i} + \frac{log\, \rho_i}{\gamma_i} t - \frac{1}{\gamma_i} log \lambda_{ts} \tag{4}$$

Averaging over all N individuals in the village and solving for $log\lambda_{ts}$ yields an expression for average village consumption $\overline{c}_{ts} \equiv \frac{1}{N}\sum_{i=1}^{N} c_{its}$, which we use to replace the common Lagrange multiplier in equation (4). We obtain:

$$c_{its} = \frac{1}{\gamma_i}\left[log\, \eta_i - \frac{\frac{1}{N}\sum_{j=1}^{N}\frac{log\eta_j}{\gamma_j}}{\frac{1}{N}\sum_{j=1}^{N}\frac{1}{\gamma_j}}\right] + \frac{1}{\gamma_i}\left[log\rho_i - \frac{\frac{1}{N}\sum_{j=1}^{N}\frac{log\rho_j}{\gamma_j}}{\frac{1}{N}\sum_{j=1}^{N}\frac{1}{\gamma_j}}\right] t + \frac{1/\gamma_i}{\frac{1}{N}\sum_{j=1}^{N}\frac{1}{\gamma_j}}\overline{c}_{ts} \tag{5}$$

[6] Assuming constant relative risk aversion (CRRA) instead yields a similar result, except that estimating equations are expressed in logs rather than levels. See, for instance, Mace (1991). The derivation is omitted here to save space.

which shows that the consumption of individual i is a linear function of the average individual consumption $\bar{c}_{ts}$ and each parameter has been suitably normalized relative to its village average. Equation (5) shows that individual i's consumption increases linarly in $1/\gamma_i$, which captures i's willingness to bear risk. More risk averse individuals consume, other thing being equal, a smaller fraction of village consumption but, thanks to the intercept, their consumption is, as we would expect, more stable. This means that individuals who are less risk averse than the rest of the village consume less in bad years, but make up for it in good years, i.e., their consumption depends more on $\bar{c}_{ts}$. We also confirm that c_{its} increases in i's relative welfare weight and relative discount factor, with the latter effect increasing over time as noted earlier.

2.2 Accounting for consumption categories

Finally, we examine the predictions that the model makes regarding specific components of consumption, e.g., food and non-food non-durables. If individuals have homothetic preferences, income elasticities are unity for all goods and consumption shares are constant. This implies that consumption of good k is simply:

$$c_{itsk} = \alpha_k c_{its}$$

In this case, model (5) applies equally to all consumption goods, except that all coefficients are premultiplied by α_k. This means that risk pooling can be tested with any component of consumption.

This is no longer the case when consumption preferences are not homothetic, i.e., when $c_{itsk} = \alpha_k(c_{its})$ where $\alpha_k(.)$ now denotes an Engel curve. If the shape of this Engle curve can be estimated separately, e.g., from an analysis of the relationship between consumption shares and total consumption expenditures in a cross-section, model (5) can still be fitted to specific consumption categories provided the dependent variable is suitable transformed as:[7]

$$\hat{c}^k_{its} \equiv \hat{\alpha}_k^{-1}(c_{itsk}) \tag{6}$$

In an efficient risk pooling economy, applying model (5) to each $\hat{c}^k_{its}$ should yield the same coefficient estimates. This would indicate that all consumption categories move with total expenditures in a way consistent with preferences across goods. It is also conceivable that risk pooling focuses more on basic necessities such as food, but ignore luxuries. In this case, the consumption share of luxuries would fall faster with a fall of total expenditures than predicted by the Engel curve, i.e., $\hat{c}^k_{its}$ would vary more with $\bar{c}_{ts}$ when k is a luxury than when k is food expenditures. This can be investigated by comparing coefficient

[7] For this transformation to yield a usable $\hat{c}^k_{its}$ in our test, function $\hat{\alpha}_k(.)$ must be monotonic over the relevant range.

estimates of model (5) applied to consumption categories $\hat{c}_{its}^k$ with low and high income elasticities.

We now summarize the main predictions of our model for efficient risk pooling:

1. Individual consumption c_{its} is independent of individual liquid assets x_{its} and individual income y_{its}.

2. Individual consumption c_{its} is a function of village aggregate consumption expenditures c_{ts}

3. Average village consumption $\overline{c}_{ts}$ is a concave function of aggregate village cash-in-hand x_{is} – i.e., the village smoothes consumption over time using all village liquid assets as pooled precautionary savings.

4. The share of village consumption that individuals receive falls over time if they are more impatient than the (suitably weighted) village average.

5. Individuals who are more risk averse than the (suitably weighted) village average receive, other things being equal, a smaller share of average village consumption. As a result their consumption is smoother than that of less risk averse individuals in the village.

6. The consumption of goods with a low income elasticity is smoothed more than the consumption of goods with a high income elasticity. Once transformed by the inverse of the Euler curve, expenditure shares on specific goods all respond identically to aggregate village expenditures c_{ts}

3 Testing strategy

Before we present our testing strategy in detail, we must first recognize that, while the model presented in Section 1 applies at the individual level, in our data, as in most, consumption consumption, assets and income are all measured at the household level. As a result, we cannot estimate the extent to which risk is pooled within households (e.g., Dercon & Krishnan (2000); Dunbar et al. (2013)). We can only test whether it is pooled across households.

To do so in a way consistent with theory, we need to normalize the data in such a way that, if risk were perfectly pooled within and across households, our methodology would conclude that it is. In order to obtain a correct village average $\overline{c}_{ts}$, we must weigh each household's per-capita consumption by the number of its members.[8] The same reasoning

[8]This is best illustrated with a simple example. Imagine two households 1 and 2, respectively with 1 and 2 members. Total consumption in household 1 is 100, which is also the consumption per head. In household 2, total consumption is also 100, which means that consumption per head is 50. If we take the

applies to income and assets, as well as to the risk sharing tests themselves. For this reason, all regressions presented in the paper are weighted by household size, so as to ensure that our tests aggregate individuals in a way that is consistent with theory.[9] In practice, we measure the size of each household by its number of adult-equivalents to reflect the fact that minimal consumption needs vary by age and gender.

3.1 Homogeneous risk and time preferences

We start by testing the predictions of the model under the assumption of homogeneous risk and time preferences. Since there is only one realized state of the world per time period, equation (5) simplifies to a perfect risk pooling relationship of the form:

$$c_{it} = \beta_i + \beta_1 t + \beta_2 \overline{c}_t \tag{7}$$

with $\beta_1 = 0$ when all ρ_i are identical. It is important to note that assets are absent from this equation. This is because, thanks to predictions 2 and 3 above, $\overline{c}_t$ is a sufficient statistic about the social planner's choice of future savings for the village. It follows that standard tests of risk pooling that use village average also work when the village saves.

As in the rest of the literature (e.g., Mace 1991, Cochrane 1991, Townsend 1994, Ravallion and XXX 1997), we first-difference equation (7) to eliminate the individual specific welfare weight term β_i. We also add two regressors, income y_{its} and assets w_{its}. This is one regressor more than the standard risk pooling test, which ignores assets and precautionary savings either at the individual or village level. The estimated CARA model has the form:

$$\Delta c_{it} = \beta_1 + \beta_2 \Delta \overline{c}_t + \beta_3 \Delta y_{it} + \beta_4 \Delta w_{it} + \epsilon_{it} \tag{8}$$

Given the relatively small number of households within each village, the mechanical correlation between c_{it} and $\overline{c}_t$ generates a bias in β_i when the null of perfect risk pooling is false (see Appendix A for illustration).[10] To correct for this bias, we estimate (10) by replacing the village mean $\overline{c}_t$ by the leave-out-mean $\overline{c}_{-i,t} \equiv \frac{1}{N-1}\sum_{j \neq i} c_{jt}$.[11] We also estimate a similar CRRA model where all variables are expressed in logs – see the Appendix for a formal derivation.

---

simple average of consumption per head across the two households we obtain average village consumption of $75 = \frac{1}{2}100 + \frac{1}{2}50$. If, however, we average across individuals, the average village consumption is $66.67 = \frac{1}{3}100 + \frac{2}{3}50$.

[9]To the best of our knowledge, however, this easy correction is not implemented by Mace (1991); Cochrane (1991); Townsend (1994), and those that followed in their footsteps (e.g., (Schulhofer-Wohl, 2011; Mazzocco & Saini, 2012; Chiappori et al., 2014)).

[10]For instance, when the true $\beta_i = 0$ in equation (10), the OLS estimate has a bias equal to $1/N$.

[11]It is easy to show that, under the null of perfect risk pooling, estimating (10) with the leave-out-mean still yields the correct estimate of $\hat{\beta}_i = 1$ but multiplies α_i by $\frac{N}{N-1}$. See Appendix A for details.

Our main null hypothesis is that risk pooling is efficient, which implies that $\beta_2 = 1$ and $\beta_1 = \beta_3 = \beta_4 = 0$. Equation (8) also enables us to consider the following alternative hypotheses:

1. Hand-to-mouth: Each individual consumes his or her income y_{its}, which implies $\beta_3 = 1$ and $\beta_1 = \beta_2 = \beta_4 = 0$
2. Individual precautionary saving: Each individual consumes a concave fraction of his or her cash-in-hand $x_{its} \equiv y_{its} + w_{its}$, which implies that $\beta_3 = \beta_4 > 0$ and $\beta_2 = 0$
3. Individual precautionary saving with excess sensitivity to income: $\beta_3 > \beta_4 > 0$ and $\beta_2 = 0$
4. Partial pooling of income but full pooling of assets: $1 > \beta_2 > 0$ and $1 > \beta_3 > 0$ and $\beta_4 = 0$
5. Partial pooling of income and assets: $1 > \beta_2 > 0$ and $\beta_3 > 0$ and $\beta_4 > 0$

This regression is complemented by a village level analysis to test whether the village collectively uses assets to smooth consumption. The estimated regression is the standard test of the precautionary saving developed by Zeldes (1989b). It takes the following form:

$$\Delta \overline{c}_t = \beta_1 + \beta_3 \Delta y_t + \beta_4 \Delta w_t + \epsilon_t \tag{9}$$

Efficient precautionary saving requires asset integration, which implies that liquid assets and income have the same effect on consumption: $\beta_3 = \beta_4$. If the village does not use assets to smooth consumption across periods, then $\beta_1 = \beta_4 = 0$ and $\beta_3 = 1$. It is also conceivable that the village achieves a modicum of intertemporal consumption smoothing from other sources that are not identified in the data (e.g., external transfers from migrants, government, or NGOs), in which case $\beta_1 > 0$ and $1 > \beta_3 \geq 0$. We also estimate (9) in log form. In addition, we present a non-parametric regression of $\Delta \overline{c}_t$ on village cash-in-hand Δx_t. We expect to find a concave relationship between consumption and cash-in-hand, as predicted by the precautionary savings model (e.g., Zeldes 1989a, Deaton 1991).

3.2 Consumption categories and Engel curves

Next we estimate inverse Engel curves (6) for various consumption goods. This is achieved by non-parametrically regressing total expenditures c_{its} on expenditures c_{itsk} on good k. We do this using cross-section data, which means that the income elasticities embedded in these inverse Engel curves are estimated using variation in expenditure shares across households with different total levels of expenditures. We then use the fitted model $\hat{\alpha}_k^{-1}(c_{itsk})$ to obtain a prediction of total expenditures $\hat{c}_{its}^k$ for each household in each period. If households are unconstrained in the consumption choices they make after risk sharing, they should, on average, be on their Engel curve for each good. In contrast, if assistance from the village favors certain goods – e.g., food[12] – then households should spend a higher proportion of their total expenditures on food when they receive assistance. This observation forms the basis of our test.[13]

To implement this idea in the simplest way, model (8) is estimated separately for each $\Delta\hat{c}_{its}^{k}$ dependent variable. We then test whether estimated coefficients β_2, β_3, and β_4 are identical across consumption goods. This constitutes an alternative test of the perfect risk pooling model. The alternative is that certain expenditures are better insured than others – e.g., luxuries are consumed when the village as a whole is enjoying a higher income and, in bad times for the village, individual consumption patterns are adjusted towards food consumption. Differentiated insurance is a common occurrence in all societies: social safety nets typically seek to guarantee individuals a minimum consumption level, with a focus on necessities such as food, shelter, and basic clothing – but typically excludes luxuries. To verify whether this pattern is also present in our data, we test whether β_2 is smaller (i.e., less sensitive to aggregate shocks) for goods with a low income elasticity, and vice-versa for goods with a high income elasticity, such as luxuries.

3.3 Heterogenous risk and time preferences

We now introduce heterogeneity in risk and time preferences across households. It is well known that tests of risk pooling are biased in the presence of heterogeneous risk preferences (Schulhofer-Wohl, 2011; Mazzocco & Saini, 2012; Chiappori et al., 2014). Ignoring heterogeneous risk preferences leads to an upwards bias in $\hat{\beta}_2$, the coefficient of $\overline{c}_t$ in equation (8). This is because $\epsilon_{it}^{homog} = \left(\frac{1}{\gamma_i} - \frac{1}{\gamma}\right)\overline{c_t} + u_{it}$, which introduces a positive correlation between $\overline{c}_t$ and the error term.

To address this issue, we proceed in two steps. We do not have (reliable) information on monthly income and assets: this information is only available annually. But we do have reliable information on monthly consumption for each households over a period of five years (60 months). We therefore have enough observations to fit a perfect risk sharing model to each household separately, since this model does not require data on income and assets. This yields estimates of household-specific risk and time preference parameters that are consistent under the null of perfect risk sharing, and thus can then be used, in

[12] as the US Food Stamps welfare program used to do.

[13] To illustrate with an example, imagine that a household optimally spends 700 Rps on food and 300 on non-food when its total expenditures is 1000 Rps, and 800 on food and 400 on non-food when its total expenditures is 1200. Then if this household is unconstrained and we observe it to spend 800 Rps on food, its total expenditures should be 1200. If, at the same time, we observe it consuming 300 on non-food, we would predict that its total expenditures is 1000. Hence a systematic discrepancy between the two predicted values of total expenditures $\hat{c}_{its}^{food}$ and $\hat{c}_{its}^{non-food}$ indicates that consumption choices are constrained.

a second step, to reestimate model (8) with household-specific risk and time preferences on annual data.

Formally, in the first step we estimate model (5) separately for each of the 1300 households in our data. To achieve this, we start by noting that, as pointed out by Wilson (1968), doubling every household's coefficient of risk aversion does not change the set of Pareto-efficient allocations. This means that absolute risk preferences cannot be identified – but relative risk preferences can. To reflect this, we follow Chiappori et al. (2014) and normalize risk preferences up to a village-specific scale by setting $\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\gamma_i} = 1$. With this normalization, equation (5) reduces to:

$$c_{it} = \frac{1}{\gamma_i}\left[\log\eta_i - \frac{1}{N}\sum_{j=1}^{N}\frac{\log\eta_j}{\gamma_j}\right] + \frac{1}{\gamma_i}\left[\log\rho_i - \frac{1}{N}\sum_{j=1}^{N}\frac{\log\rho_j}{\gamma_j}\right]t + \frac{1}{\gamma_i}\overline{c}_{-i,t}$$

We use 60 months of consumption data to estimate, for each household i, an OLS model of the form:

$$c_{it} = \alpha_i + \theta_i t + \beta_i\overline{c}_{-i,t} + \epsilon_{it} \tag{10}$$

The mapping between estimated coefficients and structural parameters is given by:

$$\beta_i = \frac{1}{\gamma_i} \tag{10A}$$

$$\alpha_i = \beta_i\left[\log\eta_i - \frac{1}{N}\sum_{j=1}^{N}\beta_j\log\eta_j\right] \tag{10B}$$

$$\theta_i = \beta_i\left[\log\rho_i - \frac{1}{N}\sum_{j=1}^{N}\beta_j\mathrm{log}\rho_j\right] \tag{10C}$$

Coefficient β_i represents the risk tolerance (i.e., the inverse of risk aversion γ_i) of individual i relative to the village mean – e.g., $\beta_i > 1$ implies that i is more risk tolerant than others in the village, and as a result has a consumption level that varies more than others with the village average.

Estimates of structural parameters γ_i, η_i, and ρ_i can be recovered from OLS estimates of $\hat{\beta}_i$, $\hat{\alpha}_i$, and $\hat{\theta}_i$, subject to suitable normalization. For time preferences, the same reasoning applies as for time preferences: only relative preferences can be recovered from (10). We therefore set $\frac{1}{N}\sum_{j=1}^{N}\rho_j = 1$. For welfare weights, we follow convention and normalize them to sum to 1 within each village. With these normalizations, $\hat{\gamma}_i = 1/\hat{\beta}_i$ and estimates of relative welfare weights η_i and relative time preference parameters ρ_i

can be recovered using the following formulas:[14]

$$\log \eta_i = \frac{\alpha_i}{\beta_i} + \log \left[\frac{1}{\sum_{j=1}^{N} e^{\frac{\alpha_j}{\beta_j}}} \right] \tag{10A}$$

$$\log \rho_i = \frac{\theta_i}{\beta_i} + \log \left[\frac{1}{\frac{1}{N} \sum_{j=1}^{N} e^{\frac{\theta_j}{\beta_j}}} \right] \tag{10B}$$

This yields a set of household-specific estimates of $\hat{\gamma}_i$, $log\,\hat{\eta}_i$, and $log\,\hat{\rho}_i$, all estimated under the maintained assumption of perfect within-village risk sharing. These estimates represent what the relative welfare weights and the relative risk and time preferences of households *would be* if income risk is perfectly shared among villagers.

The second step of our test is to use these inferred parameters to control for household-level risk and time preferences when estimating our risk pooling test with annual data on income and wealth. We extend model (8) to allow for household heterogeneity in risk and time preferences as follows:

$$c_{it} = \alpha_i + \theta_i t + \beta_i \overline{c_{-i,t}} + \xi y_{it} + \zeta w_{it} + \epsilon_{it} \tag{11}$$

As before, perfect risk pooling requires that $\xi = 0$ and $\zeta = 0$. To test this prediction, we write (11) so as to eliminate all the household-specific coefficients. First, α_i is eliminated by first-differencing the data. Second, we use the $\hat{\beta}_i$ and $\hat{\theta}_i$ estimates obtained in the first step to create an estimable model of the form:[15]

$$\Delta \left(c_{it} - \hat{\beta}_i \overline{c_t} \right) - \hat{\theta}_i = \xi \Delta y_{it} + \zeta \Delta w_{it} + \Delta \epsilon_{it} \tag{12}$$

4 Data

We use the new wave of ICRISAT's VDSA (Village Dynamics of South Asia) panel data of about 1400 households observed over 60 consecutive months from June 2010 to

[14]Let $X = -\frac{1}{N} \sum\limits_{j=1}^{N} \beta_j \log \eta_j$ and $\phi_i = \frac{\alpha_i}{\beta_i}$. From equation (10C) we get $\log \eta_i = \phi_i + X$ (*) and thus $\eta_i = \exp^{\phi_i + X}$. By the normalization of welfare weights $\sum_{j=1}^{N} \eta_j = 1$ we get $\sum_{j=1}^{N} e^{\phi_j + X} = 1$, which implies $X = \log \left[\frac{1}{\sum_{j=1}^{N} e^{\phi_j}} \right]$. Substituting X back into (*) yields the reported formula. A similar approach yields the ρ_i formula, except for the division by N which comes from the different normalization rule $\frac{1}{N} \sum\limits_{j=1}^{N} \rho_j = 1$.

[15]Although regression model (12) makes use of predicted variables $\hat{\beta}_i \overline{c_t}$ and $\hat{\theta}_i$, estimates of ξ and ζ are not subject to sampling error since the contructed variables only appear in the dependent variable (e.g., Murphy & Topel (1985)).

July 2015.[16] Households were randomly selected from 30 villages in eight Indian states, chosen to represent the agro-climatic conditions in India's semi-arid and humid tropical regions.[17] Households in each village were randomly selected to represent households in four landholding classes: large, medium, small, and landless. The data collection timeline follows the agricultural cycle in India, beginning from June to July. Attrition in the VDSA data is minimal - only about 10% of households have an unbalanced panel of less than 60 months of data. For our analysis, we use a balanced panel of 1,296 households that reported 60 months of consecutive monthly data.

To construct the main consumption outcomes, we use data on food expenditures, non-food expenditures and total expenditures collected every month for each household. Food consumption includes all food items sourced from home production and purchases. Non-food consumption includes expenses on services and utilities such as travel, education, medical, and energy.

Our measure of earned income includes all net earnings from crops, livestock and off-farm labor. Crop and livestock income is calculated as the revenue from sales of crop and livestock products, minus production costs that include the value of material inputs and the imputed cost of own labor. Off-farm labor income is the sum of earned wages for all household members and the net income earned from household businesses. The majority of individuals in the sample are at least partially employed in the casual labor market. A few individuals are employed in business or a salaried job in the formal sector.

In the analysis we use two measure of household assets: liquid wealth and cash-in-hand. Liquid wealth is defined as the sum of the household's net credit position (savings, minus borrowing plus lending) and the value of liquid assets such as livestock, consumer durables, and inventories of crops, inputs, and fueld. Cash-in-hand is constructed as the sum of liquid wealth and earned income.

Although the VDSA has rich monthly data on consumption and income, household assets are only measured annually at the beginning of each panel year, which coincides with the onset of the main agricultural season in June. Consequently, all regressions that require asset information are estimated by aggregating monthly data on household consumption and income to the beginning of the agricultural cycle. All values are deflated and expressed in 2010 Indian rupees. Income, consumption, and assets are expressed per

[16] ICRISAT's new wave of VDSA panel data is a continuation of Village level studies (VLS) panel of household data collected between 1975 to 1985 in six villages in the semi-arid tropics of India. In the VDSA data, in addition to the 6 old VLS villages, 12 more villages in the semi-arid tropics and 10 more villages from East India were included, summing to a total of 30 villages across 8 states in India. The VDSA data collection started in 2009, however, the data for panel year 2009 has many gaps, especially in the consumption module. To maintain consistency, this paper uses data beginning from panel year 2010 until 2014.

[17] The eight states are Andhra Pradesh, Bihar, Gujarat, Jharkhand, Karnataka,Madhya Pradesh, Maharashtra, and Orissa). Four villages were selected from each state, except in Madhya Pradesh where only two were selected. See figure A1 in Appendix for the precise location of the 30 villages.

capita by dividing them by their adult-equivalent weight.[18]. We also trim the top and bottom 1% of the data to remove outliers and large measurement errors.[19]

Table 1 presents descriptive statistics for the main variables used in the analysis. Annual consumption expenditures per adult equivalent are on average Rs. 14,961 in 2010 rupees. This is equivalent to 2.89 US$ per day and per adult-equivalent, based on a purchasing parity rate of 14.59 Rs. per US$ in 2010 (World Bank, 2014).

5 Preparatory analysis

5.1 Engel curves

Before launching the main part of our analysis, we complete the preparatory analysis on Engel curves and heterogeneous preferences. We begin by fitting Engel curves to annual consumption data in 2011, a good rainfall year when village cash-in-hand is the highest in our data. This is the year in which risk pooling would be least likely to impose constraints on consumption. Figure 1 uses a flexible polynomial to plot household budget shares against the log of total household expenditure per capita. The Engel curve for food is approximately log-linear and downward sloping, confirming that the food expenditure share falls with income, in accordance with the literature. The poorest quintile of the income distribution spend about 65% of their total budget on food, whereas the richest quintile spend about 40%. The inverse is true for non-food expenditures. These results constitute strong evidence against homotheticity in food and non-food preferences. Since the computed Engel curves are monotonic, they can be inverted to obtain the function $\hat{\alpha}_k^{-1}(c_{itsk})$.

5.2 Risk and time preferences

Next, we estimate individual risk and time preferences as described in Section 3.3. To recall, we estimate the CARA-based regression (10) for each household, only using monthly data on consumption expenditures. We then use the results to recover estimates of absolute risk aversion and time preference by using the formulas reported in Section 3.3. The risk tolerance measure β_i is normalized to a village-specific scale – i.e., mean risk tolerance of each village is set to one. To recall, risk tolerance is the inverse of the coefficient of risk aversion. The estimate of the discount factor ρ_i is similarly normalized to average to one in each village. These normalizations arise from the fact that only relative values of risk aversion and time preference can be inferred from the coefficients

[18] Following Townsend (1994), we define the age-sex weights as : 1.0 and 0.9 for adult males and females; 0.94 and 0.83 for adolescent males and females aged 13-18, 0.67 for children aged 7-12 regardless of gender; 0.52 for toddlers 1-3 and 0.05 for infants

[19] For assets and wealth, we lose about 130 observations out of a total of 6500 annual observations.

of regression (10). We also estimate a CRRA version of these parameters using a model similar to (10), but in logs. Apart from the normalizations, it is important to remember that the estimated risk and time preference parameters are obtained under the *maintained assumption* of perfect risk pooling and that their main purpose is to test perfect risk pooling in the heterogeneous-corrected regression model (12). This being said, these estimated parameters contain valuable information that we summarize here.

In Figure 2 we plot the distributions of the household-specific estimates of risk tolerance $\hat{\beta}_i$ under both the CARA and CRRA models.[20] These parameters are identified from whether household i's consumption varies more than that of household j: if it does – and we are in a perfect risk pooling equilibrium – then i must be less risk averse than j. Both sets of estimates are normalized to have a mean equal to 1 within each village, which means that they capture relative risk tolerance rather than absolute values. Since the CARA and CRRA estimates are not measured in the same units, their magnitude is not directly comparable; but their frequency distribution is.

Overall, we find considerable heterogeneity in risk tolerance within villages, suggesting that, if we are in a perfect risk pooling equilibrium, large welfare benefits are achieved not only from pooling risk, but also from shifting risk from highly risk averse households to more risk neutral ones. Since more variation in β_i around its average $\overline{\beta}$ translates into more correlation between $\overline{c_t}$ and u_{it} in $\epsilon_{it}^{homog} = \left(\beta_i - \overline{\beta}\right) \overline{c_t} + u_{it}$, Figure 2 also constitues prima facie evidence that ignoring heterogeneity in risk preferences may bias risk pooling tests that assume homogeneous risk preferences.

Figure 3 shows the distribution of household-specific estimates of discount factors $\hat{\rho}_i$. As for risk preferences, the $\hat{\rho}_i$ are normalized to have unit mean within each village. Discount factors are identified from whether household i's consumption increases more over time than that of household j: if it does – and we are in a perfect risk pooling equilibrium – then i must be more patient than j. We see that, visually, there is much less disperson in discount factors than in risk tolerance. This suggests that, if the study area is in a perfect risk pooling equilibrium, few welfare gains are achieved by accommodating differences in impatience. This being said, even small differences in discount factors can, over time, translate into increasing differences in consumption levels across households, even if they share the same risk preferences and the same welfare weight.

[20]Because each $\hat{\beta}_i$ is estimated from a regression with 60 observations, its distribution suffers from excess variance due to sampling error. To assess the magnitude of the excess variation that this error generates in Figures 2 and 3, we shrunk the sample distribution of $\hat{\beta}_i$ in such a way that, when we add the sampling noise to the 'shrunk' $\hat{\beta}_i$, we obtain a frequency distribution with the same variance as that of the original $\hat{\beta}_i$. In practice, this procedure entails turning the original $\hat{\beta}_i$ distribution to have zero mean, and using the standard error of $\hat{\beta}_i$ in each regression as estimate of the sampling error in that sample. Using this approach shrinks the standard deviation of the estimated parameters by 17% on average, while respecting the general shape of the original distribution. The qualitative conclusions reported here are not affected, however.

5.3 Inequality

We also infer welfare weights $\hat{\eta}_i$ from estimated coefficients from regression (10), using the formula presented in Section 3.3. The welfare weights are identified from the household-specific intercept in model (10). In the CARA model, this intercept measures the minimum level of consumption that is assigned at $t = 0$ to household i in a perfect risk pooling equilibrium. Keeping risk and time preferences constant, and keeping $\overline{c}_t$ the same, household i consumes more that j if i has a higher $\hat{\eta}_i$ than j. In the CRRA version, the intercept is the base *share* of consumption that goes to i, but the reasoning is the same: ceteris paribus, i consumes more if i has a larger welfare weight. Following common practice, the welfare weights themselves are normalized to sum to 1 within each village. It follows that equal treatment of all households in a village of size N_v requires that they all have $\hat{\eta}_i = 1/N_v$. Since N_v varies across villages, it is useful to take $1/N_v$ as yardstick to judge intra-village inequality.

Using this approach, we calculate, for each village, the proportion of households for whom $\hat{\eta}_i < 1/N_v$. The larger this proportion is, the more unequal the distribution of welfare is in the village. We present in Figure 4 a histogram of these proportion across all the villages in our study. While there is some variation between the histograms depending on whether the welfare weights were estimated using CARA or CRRA, it is nonetheless clear that welfare weights are quite unequal in most villages. Across all villages, the proportion of households with welfare weights less than the equitable share $1/N_v$ is estimated to be about 88% from the CARA version of regression (10), and 89% for the CRRA version.

Figure 4 presents two histograms of village-averages of the 30 villages, one from the CARA $\hat{\eta}_i$ estimates and the other from the CRRA estimates. In most of the villages, more than 95% of households have welfare weights less than $1/N_v$, and the overwhelming majority of them have 80% or more of households below the average welfare weight of $1/N_v$. This implies that these villages have 20% or less of their households enjoying above average welfare weights – and thus consistently above average consumption across time. We also find that the frequency distribution of $\hat{\eta}_i$ has a fat upper tail, with some households receiving welfare weights close 1, indicative of very high consumption inequality. Keeping in mind that these estimates all assume perfect risk pooling at the village level, they remind us that risk pooling is not equivalent to income redistribution – and that it is quite compatible with a lot of consumption inequality in equilibrium. This inequality would be further reinforced in good years if richer households – i.e., those with a high welfare weight and, thus, a high average consumption – are also those who are more risk neutral, as is likely.

6 Village-level analysis

We now turn to our main estimation results. We start by reporting the results of the risk pooling tests under the assumption of CARA utility. We then repeat the exercise for the CRRA model to check the robustness of our findings. Next we estimate the extent of precautionary saving at the village level. In the last part of this Section, we reestimate our main results for perfect risk pooling within castes (Jatis) in the same villages, instead of within villages.

6.1 CARA model

Table 2 summarizes all our test results for perfect risk pooling within villages under a CARA model. As explained in Section **??**, these estimates are obtained from first-difference regressions in levels, using a pooled panel of all the sample households. In all regression, standard errors are clustered at the village-level. Panel A in Table 2 reports the test results under homogeneous preferences. As shown in Column (1) Panel A, for total expenditures, full risk pooling of income and assets is rejected. The small magnitude of the co-efficients on income and wealth nonetheless suggest substantial risk pooling. A Rs.100 change in annual income is associated with Rs. 3.14 change in total annual consumption, all measured in real 2010 rupees per-adult equivalent. Similarly, a Rs.100 change in annual liquid wealth is associated with Rs. 1.08 change in annual consumption. The coefficient on village expenditure (0.923) is not different from 1 at the 5% significance level.[21] This indicates a high degree of mutual risk pooling within villages, thereby rejecting the pure hand-to-mouth or individual precautionary savings models discussed in Section 3.1. We, however, note an excess sensitivity of consumption to household income and assets: the coefficients of income and assets are both statistically significant. The absolute value of these coefficients, however, is well below 1, suggesting partial pooling. Taken together, these findings are consistent with a pooling of income and assets that is partial but nonetheless achieves a considerable amount of co-movement in household consumption across years. We can also reject a model in which assets are pooled for risk purposes but incomes are not.

Next, we examine the implications of risk pooling separately for food and non-food, under homothetic and non-homothetic preferences. Under homethetic preferences, consumption shares are assumed constant with income, which implies that, in the absence of constraints on consumption, each consumption expenditure category should, on average, respond equally to village average expenditures as well as income and asset shocks. If, in contrast, food consumption is better insured than non-food consumption, food ex-

[21] In Appendix B we offer a simple back-of-the-envelope correction for the downward bias due to sampling error in $\bar{c}_{-i,t}$. Applied to this coefficient, the correction yields a point estimate of $0.9236/0.976 = 0.9463$, a figure that is not even significantly different from 1 at the 10% level.

penditures should respond more to village average expenditures and less to variation in household assets and income. Results under the assumption of homothetic preferences are shown in columns (2) and (3) for food and non-food expenditures, respectively. To maintain direct comparability for column (1), food and non-food expenditures are divided by the *average* budget share for food and non-food, respectively. For instance, if the average food share is 50%, the food expenditure variable is multiplied by 2, making it comparable to the total expenditure variable used in column (1). With this in mind, we see that, if anything, non-food expenditures respond more to average village expenditures. We also note that non-food expenditures respond more strongly to variations in household liquid wealth and income, suggesting less smoothing of non-food expenditures across individuals. Taken together, these findings suggest that year-to-year variation in household expenditures on food and non-food depart from what cross-section expenditure shares would suggest. This indicates that food expenditures are not only less responsive to income and assets shocks – i.e., they are better protected from idiosyncratic shocks – but also that they fluctuate less with average village consumption than non-food expenditures – i.e., they are better protected from collective shocks. The contrast between food and non-food expenditures suggests that this is achieved by smoothing food consumption from collective shocks than non-food expenditures.

This interpretation, however, can be misleading because it incorrectly assumes that consumption preferences are homothetic. Under non-homothetic preferences, the approach must be amended to account for the systematic variation of expenditure shares with total expenditures. To correct for this, we transform the dependent variable by the inverse of the Engel curve (see Section 3.2 for details). As a result, it becomes the level of total expenditures that is predicted from observed food expenditures and cross-section Engel curve estimates for year 2011. To illustrate with an example, let the food expenditures per capita of household i be Rs.1000. Further suppose that, based on the Engel curve, a household with a total expenditures of Rs.2500 has an average food share of 0.4 and thus spends Rs.1000 on food. It follows that a household that spends Rs.1000 on food must, on average, have a total expenditure of Rs.2500 in order to be on its Engel curve. Using this approach, we can test whether variation in food and non-food expenditures responds similarly to village average expenditures and income and asset fluctuation – as they should if total consumption is redistributed among households but households can spend optimally. A side benefit of this transformation is that the dependent variables in columns (4) and (5) are expressed in the same units as in columns (1) to (3), making coefficients comparable between them.

As anticipated, we find that the β_2 coefficient estimates shown in columns (4) and (5) are less different from each other than the coefficient estimates obtained by assuming homothetic preferences. This confirms that assuming homothetic preferences over-estimates the excess sensitivity of non-food expenditures to village shocks relative to food expenditures: part of this higher sensitivity is due to the fact that non-foods have a higher income elasticity. With this correction, we nonetheless continue to observe that non-foods respond more to aggregate village shocks than food expenditures.

Taken together, these findings indicate that food expenditures (which have low income elasticity) are more insulated from from village-level shocks than non-food expenditures (which have high income elasticity) *over and beyond* the smoothing in food consumption that would arise purely as a result of optimal reorganization of consumption towards food when individual total expenditures fall. The converse is true for non-food expenditures since they are, on average, more volatile than would be optimal based on their cross-section income elasticity.

In Panel B of Table 2, we reestimate all regressions while allowing for heterogeneous risk and time preferences. As described in Section 3.3, this involves two steps. In the first step, we estimate *individual* households' risk and time preferences using 60 months of consumption data and assuming perfect risk pooling. The results from this estimation were discussed in Section 5. The second step relies on equation (12) to estimate the coefficients of liquid wealth and earned income. As shown in Column (1) of Panel B for total expenditures, correcting for the bias from heterogeneous preferences does, as anticipated, reduce the magnitude of the coefficients on income and liquid wealth. But the difference is minimal, suggesting that the bias is small. A similar conclusion emerges for columns (2) to (5): coefficient estimates in Panel B are similar to those reported in Panel A. Overall, this confirms our earlier interpretation of the findings.

6.2 CRRA model

Next we re-estimate all the regressions presented in Table 2 under the CRRA assumption. The main change is that the dependent variable and the regressors are now expressed in (first differences of) logs instead of levels. Other changes relate to the way risk and time preferences are estimated – a point already discussed in Section 5. The interpretation of the reported coefficients nonetheless remains the same.

The results are presented in Table 3. A number of observations are lost when taking logs due to zero or negative values in liquid wealth or earned income. Negative values arise, for instance, when a household is a net borrower or when the imputed value of inputs (including family labor) allocated to crops and livestock exceeds crop and livestock revenues –e.g., due to crop or livestock losses. This loss of observations means we suffer some loss of power compared to Table 2. Standard errors are clustered at the village-level in all regressions.

The first thing to notice is how similar results in Panel A are to those reported in Table 2. The coefficient of average village expenditures in column (1) is significantly below 1[22]

[22]The coefficient of $\bar{c}_t$ remains significantly different from 1 at the 1% level even when we apply the

– indicating less than perfect risk pooling – but the magnitude of the difference is not large. Similarly, we find that year-to-year variations in household-level liquid wealth and earned income have a statistically significant effect on consumption expenditures – but the magnitude of the effect is very small: a doubling of income, for instance, translates, on average, into a 1.8% increase in household consumption per capita; and a doubling of liquid wealth into a 2.6% increase in consumption.

The first estimate is interpretable as the marginal propensity to consume (MPC) out of income and is precisely estimated with a standard error of 0.006. It is slightly lower than the MPC of 0.05 estimate by Blundell et al. (2008) for households in US using PSID data, and is on the lower end of the mean MPC of 0.21 reported in a recent meta-analysis of 246 MPC estimates by Havranek & Sokolova (2020). The low MPC found in our study is far lower from the original estimate of 0.5 by Campbell & Mankiw (1989). It implies that a very low proportion of households in the VDSA villages live 'hand-to'mouth' and that most households engage in consumption smoothing strategies of mutual risk pooling.

Turning to Table 3 columns (1) to (5), we also find that food and non-food expenditures do not move with total expenditures in the same way across years within households as they do across households within years. If we take variation across households within years to compute average consumption shares (columns 2 and 3) or Engel curves (columns 4 and 5), we again find that correcting for non-homogeneity in consumption reduces the difference in estimated sensitivity to village shocks in average expenditures. We nonetheless find that, even with the correction for non-homogeneity, food and non-food still do not vary with average village consumption in a way consistent with a within-year utility-maximizing allocation of total expenditures: households increase their food consumption less – and their non-food consumption more – than what the increase in average village expenditures would predict.

We also do not find that non-food expenditures respond more to household income and wealth than food expenditures, ruling out the idea that there is less risk pooling in non-food consumption than in food consumption. Instead, the results suggest that, in a good year for the village as a whole, people increase their non-food expenditures more than implied by their cross-section income elasticity, thereby sheltering their average food consumption from aggregate shocks. Why this is the case is not a question that our estimation is designed to answer, but it could be due to social pressure to avoid spending on luxuries when other villagers are hungry.

Lastly, Panel B in Table 3 presents the test results of full risk pooling after accounting for heterogeneous risk and time preferences. In Column (1) we find a large reduction the coefficients of liquid wealth and earned income in Panel B compared to Panel A. This suggests that, in the case of the CRRA model, correcting for heterogeneous preferences does completely eliminate the correlation between household consumption and household

correction for sampling error outlined in Appendix B.

income and wealth. Once we account for under heterogeneous preferences, we fail to reject full risk pooling – which is a remarkable result. This finding extends to food and non-food expenditures, whether we assume homothetic consumption preferences or allow for non-unitary income elasticities of consumption.

Overall, the results under CRRA model indicate that accounting for heterogeneity in risk and time preferences and non-unitary income elasticities are important when testing for risk pooling. The data indicate that food consumption is smoothed across years to a greater degree than non-food consumption. This could arise naturally from the fact the income elasticity of non-food expenditures is larger than for food expenditures. Results in Panel A and B demonstrate that this is not the case: the difference in the variation of food and non-food expenditure exceeds what can be predicted from the variation of expenditure shares with total consumption. Taken together, these findings indicate that food consumption is better insured than non-food consumption against village-level risk. These results suggest that village-level risk pooling aims to smooth year-to-year food consumption more than non-food consumption. This consistent with our initial hypothesis that risk pooling institutional arrangements put more emphasis on necessities such as food – and that this is achieved, at least in part, by preventing or discouraging non-food expenditures in bad years, while allowing their explosion in good years.

6.3 Village precautionary savings

So far we have shown that variation in household consumption expenditures is strongly correlated with changes in village consumption expenditures, and only mildly correlated with variation in household annual income or liquid assets. This suggests that the sharing of idiosyncratic income and asset risk is close to optimal.

We now examine whether our study villages are also close to optimal in terms of smoothing aggregate shocks. To this effect, we investigate whether the village pools precautionary savings to partially or fully insulate average village consumption from income and asset shocks. This is done using model 9 to estimate the response of total village consumption to village cash-in-hand. Efficient precautionary saving predicts a positive and significant response of consumption to cash-in-hand: when available cash falls, consumption is reduced, but less than one-for-one. For a large enough liquid wealth, consumption asymptotes certainty equivalent consumption whereby the response of consumption to an income shock is equal to the discount rate (e.g., Zeldes (1989)). To illustrate, if the discount rate is 5%, certainty equivalent consumption changes by approximately Rs.5 in response to a Rs.100 temporary increase in income – and the coefficient of cash-in-hand in regression 9 should be around 0.05. At the same time, we should also observe a stable aggregate consumption over time and a large stock of liquid wealth.

Model 9 also includes village income as regressor to estimate the excess sensitivity of (aggregate) consumption to current income. Under the null of efficient savings at the village level, there should be no excess sensitivity to income since it is already included in cash-in-hand, and the coefficient of income should be zero.

Estimated results are reported in Table 4. Panel A presents the coefficient estimates for a model in first differences (CARA) and Panel B does the same for a model in log differences (CRRA). The coefficient of cash-in-hand Column (1) of Panel A shows a small but significant response of village consumption to village cash-in-hand. Based on this coefficient estimate, a Rs.100 change in village cash-in-hand is associated with Rs. 3.1 change in total village consumption – a ratio (3%) equal to (or possibly lower than) what we would expect the discount rate to be for the study population. This at prima facie indicates a high level of consumption smoothing through precautionary saving – i.e., close to certainty equivalence. This is, however, partially negated by the larger coefficient on village earned income. This excess sensitivity to income suggests that villages do not optimally draw on liquid assets to smooth aggregate income shocks.

Columns (2) to (5) repeat the same procedure on food and non-food expenditures separately. As before, columns (2) and (3) assume linear Engel curves, that is, constant expenditure shares, while columns (4) and (5) allow expenditure shares to vary systematically with total expenditures. Results indicate that village food consumption responds little to cash-in-hand while village non-food expenditures respond significantly to cash-in-hand – and that this difference is somewhat muted when correcting for non-homotheticity. This is consistent with earlier results. Both forms of consumption display excess income sensitivity, however. Moreover, once we correct for non-homotheticity, we find that food expenditures respond more to village income than non-foods, although the difference in coefficient estimates is not statistically significant.

The CRRA results shown in Panel B are broadly similar to those reported in Panel A for a model in first differences. For sensitivity to cash-in-hand, we obtain coefficients of a similar magnitude although not statistically significant. The coefficient on earned income remains large in magnitude, confirming the excess sensitivity of village expenditures to annual income. This is also true for both food and non-food expenditures, albeit not statistically significant for non-foods, possibly due to lower power.[23]

Taken together, these findings appear to suggest that study villages achieve a high level of across-year consumption smoothing from precautionary savings, even if the perfect integration of income and liquid wealth is not achieved. This interpretation is, however, partially contradicted by the large unexplained fluctuations in village consumption implied by the low reported R^2 in Table 4. These results suggest the existence of large changes in average village consumption that do not closely follow the changes in reported village income and liquid wealth. To investigate the magnitude of this issue, we calculate

[23] Indeed, the Table shows that the R^2 for non-foods is less than half that for foods, suggesting a higher variance in non-food expenditures across villages than for foods.

the unexplained gap G_{vt} between village-level consumption and cash-in-hand for village v in year t, which we calculate as:

$$G_{vt} \equiv C_{vt} - Y_{vt} - L_{v,t-1} + L_{vt}$$

where C_{vt} is consumption in year t, Y_{vt} is income in year t, and L_{vt} is the liquid wealth at the end of year t. In principle, G_{vt} should be zero or close to zero in a precautionary savings model where consumption is financed out of income and year-to-year changes in liquid wealth. We indeed find that the mean of G_{vt} is close to zero – see Appendix Figure A2. But its variance is about four times the variance of C_{vt}.[24]

This indicates the presence of variation in village consumption that is driven by sources other than variation in earned income Y_{vt} and liquid wealth and indebtedness $L_{vt} - L_{v,t-1}$. What are these other sources of fluctuations is unclear. Measurement error undoubtedly plays a part – income is notoriously difficult to measure in rural economies and consumption expenditure data is subject to recall bias. But village averaging should, in principle, smooth out some of these measurement errors and diminish their impact on the estimation. Other possibilities include fluctuations in remittance flows from migrant family members, variation in transfers to and from relatives in other villages, and unreported financial transactions – such as unrecorded debt forgiveness. This calls for further research.

7 Caste-level analysis

So far we have focused on risk pooling within entire villages. A number of studies have however shown that, in the context of India, villages may not be an appropriate unit of risk pooling: endogamous marriage groups called *Jati* or sub-castes may be a more likely social unit within which income sharing takes place, if only because of the strong bonds they create through marriage and family-based social events (e.g., (Townsend, 1994) (Mazzocco & Saini, 2012; Shrinivas & Fafchamps, 2018)). In this section, we replicate our analysis at the level of sub-caste units within study villages.[25]

The testing strategy is identical to that used in Section 6, except that village units are replaced by sub-castes. In particular, for the regressions with homogeneous preferences, we use average sub-caste expenditures instead of village expenditures. For the regressions with heterogeneous preferences, we first estimate risk and time preferences at the individual household level, using monthly data as before. We then normalize household

[24]The variance of village consumption C_{vt} is 116,672 while the variance of the gap G_{vt} is 502,025. A test that the ratio of the two variances is equal to 1 is rejected with a p-value= 0.000.

[25]To identify caste groups, we use the available VDSA data on *Jati*/sub-caste of each household. There are a total 251 unique village-specific sub-caste groups in the data. For the analysis, we drop the sub-castes counting only one household and focus on the 168 sub-caste groups with at least 2 households.

preference estimates to the level of their sub-caste.

Table 5 contains the results for both the CARA and the CRRA model is thus the equivalent of Tables 2 and 3. Contrary to expectations based on the literature, the sub-caste results are, if anything, less compelling than those at the village level. In particular, all estimated coefficients for the average sub-caste expenditure variable are smaller – and thus more different from one – than in the regressions using village expenditures. A similar pattern can be seen for liquid wealth and earned income coefficients which tend to be slightly larger. Taken together, this suggests that risk pooling within sub-castes with each village is less strong than pooling across all households in the village.[26] This being said, the two sets of results are qualitatively similar in terms of patterns across regression models. This provides additional support for our earlier conclusions regarding the difference between food and non-food consumption smoothing and the stronger results using the model with a correction for non-homothetic preferences. It also confirms that the conclusions drawn in Panel A are not an artifact of ignoring the heterogeneity of risk and time preferences across households: conclusions are similar in sign and significance with those in Panel B. The results do, nonetheless, confirm that imposing homogeneity of preferences generates a bias: the coefficients on household income and liquid wealth get slightly smaller in Panel B – but not so much smaller as to change the take-away message of the analysis.

Table 6 replicates the precautionary saving regression analysis presented in Table 4, at the level of sub-castes. The main change is a drastic increase in the number observations, a change that increases power across the board. The pattern of results is similar to that discussed in Table 4, except that coefficient estimates are slightly smaller in magnitude. Conclusions are similar: sensitivity to cash-in-hand is very small – probably smaller than it should be given the relatively low level of liquid wealth in the data. Indeed, in a poor population such as the one we study, we would expect a stronger dependence of consumption to cash-in-hand in a precautionary savings optimum. This concern is somewhat confirmed by noting that, if anything, the unexplained variation in group-level consumption is larger in Table 6 than in Table 4.[27] We again note excess sensitivity to income, as in Table 4, except that the estimated coefficients are smaller, especially in Panel B. This suggest that, at the time our data were collected, sub-castes were not the social and economic unit at which risk pooling was taking place.

[26] See Appendix B where we show that sampling error in $\bar{c}_{-i,t}$ is not large enough to explain the difference of β_2 from 1.

[27] The R^2 in column 1 of Table 4 is 0.074 in Panel A and 0.059 in Panel B, compared to 0.053 and 0.021 in Table 6, respectively.

8 Conclusion

In this paper we have revisited the seminal and influential literature on risk pooling in village economies. We make a number of methodological contributions. First, we show how precautionary savings and non-homothetic preferences can be incorporated in standard tests of risk pooling in a simple and easy-to-implement way. Second, we expand the recent work of (Schulhofer-Wohl, 2011; Mazzocco & Saini, 2012; Chiappori et al., 2014) on heterogeneity in risk preferences to also allow for individual heterogeneity in time preferences. Third, we integrate all these improvements into a single, elegant testing strategy using more recent data.

The usefulness of our proposed approach is illustrated in a geographical and cultural context similar to that studied by Townsend (1994). While our results on aggregate risk pooling mirror those of the existing literature, our approach generates a number of new insights. First, we show that risk pooling creates a distortion in consumption such that food consumption is better protected from aggregate village shocks than non-food consumption, even accounting for non-unitary income elasticities. This finding echoes many social welfare policies in developed economies, which similarly prioretize specific types of consumption by the poor (e.g., food stamps, housing and shelter, health care, primary and secondary education, public transport) while taxing luxuries.

Second, we find that, contrary to findings based on earlier data, risk pooling in Indian villages no longer appears to occur primarily at the sub-caste level rather than at the village level; household consumption better tracks average village expenditures than average sub-caste expenditures within villages.

Third, we find evidence that household consumption tracks aggregate village cash-in-hand, suggesting some form of precautionary savings at the village level. But there is considerable excess sensitivity to aggregate income, indicating a lack of full asset integration. We also find a large unexplained gap between the variation in measured consumption expenditures and cash-in-hand at the aggregate village level. This unexplained gap is probably due, in part, to mismeasurement in household expenditures, income, and wealth. But village averaging should, in principle, smooth out some of these measurement errors. Other possibilities include fluctuations in remittance flows from migrant family members, variation in transfers to and from relatives in other villages, and unreported financial transactions – such as unrecorded debt forgiveness. More research is needed on these possible sources of insurance against village-level collective shocks.

As final observation, we should make clear that the presence of risk pooling in a village does not, by itself, constitute evidence that villagers explicitly share risk in the form of mutual insurance arrangements or self-help groups, or via contingent transfers and gift exchange. Within village risk pooling can indeed be achieved by other means, such as informal peer-to-peer loans, precautionary savings, and credit from MFIs and shopkeepers.

References

Aguiar, M. A., Bils, M., & Boar, C. (2020). *Who Are the Hand-to-Mouth?* Working Paper 26643, National Bureau of Economic Research.

Almås, I., Beatty, T. K., & Crossley, T. F. (2018). *Lost in Translation: What do Engel Curves Tell us about the Cost of Living?* CESifo Working Paper Series 6886, CESifo.

Altonji, J. G., Hayashi, F., & Kotlikoff, L. J. (1992). Is the extended family altruistically linked? direct tests using micro data. *The American Economic Review*, 82(5), 1177–1198.

Ambrus, A. & Elliott, M. (2020). Investments in social ties, risk sharing, and inequality. *The Review of Economic Studies*, 88(4), 1624–1664.

Ambrus, A., Gao, W., & Milan, P. (2021). Informal Risk Sharing with Local Information. *The Review of Economic Studies*. rdab091.

Ambrus, A., Mobius, M., & Szeidl, A. (2014). Consumption risk-sharing in social networks. *American Economic Review*, 104(1), 149–82.

Atkin, D., Faber, B., Fally, T., & Gonzalez-Navarro, M. (2020). *Measuring Welfare and Inequality with Incomplete Price Information*. Working Paper 26890, National Bureau of Economic Research.

Blundell, R., Pistaferri, L., & Preston, I. (2008). Consumption inequality and partial insurance. *American Economic Review*, 98(5), 1887–1921.

Campbell, J. Y. & Mankiw, N. G. (1989). *Consumption, Income, and Interest Rates: Reinterpreting the Time Series Evidence*. Working Paper 2924, National Bureau of Economic Research.

Chandrasekhar, A. G., Townsend, R., & Xandri, J. P. (2018). *Financial Centrality and Liquidity Provision*. Working Paper 24406, National Bureau of Economic Research.

Chiappori, P.-A., Samphantharak, K., Schulhofer-Wohl, S., & Townsend, R. M. (2014). Heterogeneity and risk sharing in village economies. *Quantitative Economics*, 5(1), 1–27.

Coate, S. & Ravallion, M. (1993). Reciprocity without commitment: Characterization and performance of informal insurance arrangements. *Journal of Development Economics*, 40(1), 1 – 24.

Cochrane, J. H. (1991). A simple test of consumption insurance. *Journal of Political Economy*, 99(5), 957–976.

De Weerdt, J. & Dercon, S. (2006). Risk-sharing networks and insurance against illness. *Journal of Development Economics*, 81(2), 337 – 356.

Deaton, A. (1991). Saving and liquidity constraints. *Econometrica*, 59(5), 1221–1248.

Deaton, A. (1992). Saving and Income Smoothing in Cote d'Ivoire. *Journal of African Economies*, 1(1), 1–24.

Dercon, S., De Weerdt, J., Bold, T., & Pankhurst, A. (2006). Group-based funeral insurance in ethiopia and tanzania. *World Development*, 34(4), 685–703.

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

Dubois, P. (2000). Assurance compléte, hétérogénéité des préférences et métayage au Pakistan. *Annales d'Économie et de Statistique*, (59), 1–37.

Dunbar, G. R., Lewbel, A., & Pendakur, K. (2013). Children's resources in collective households: Identification, estimation, and an application to child poverty in malawi. *American Economic Review*, 103(1), 438–71.

Escanciano, J. C., Hoderlein, S., Lewbel, A., Linton, O., & Srisuma, S. (2021). Nonparametric Euler Equation Identification And Estimation. *Econometric Theory*, 37(5), 851–891.

Fafchamps, M. & Lund, S. (2003). Risk-sharing networks in rural philippines. *Journal of Development Economics*, 71(2), 261–287.

Fafchamps, M., Udry, C., & Czukas, K. (1998). Drought and saving in west africa: are livestock a buffer stock? *Journal of Development Economics*, 55(2), 273–305.

Gertler, P. & Gruber, J. (2002). Insuring consumption against illness. *American Economic Review*, 92(1), 51–70.

Golosov, M., Kocherlakota, N., & Tsyvinski, A. (2003). Optimal indirect and capital taxation. *The Review of Economic Studies*, 70(3), 569–587.

Hastings, J. & Shapiro, J. M. (2018). How are snap benefits spent? evidence from a retail panel. *American Economic Review*, 108(12), 3493–3540.

Havranek, T. & Sokolova, A. (2020). Do consumers really follow a rule of thumb? three thousand estimates from 144 studies say "probably not". *Review of Economic Dynamics*, 35, 97–122.

Kazianga, H. & Udry, C. (2006). Consumption smoothing? livestock, insurance and drought in rural burkina faso. *Journal of Development Economics*, 79(2), 413–446. Special Issue in honor of Pranab Bardhan.

Kimball, M. S. (1988). Farmers' cooperatives as behavior toward risk. *The American Economic Review*, 78(1), 224–232.

Kinnan, C. (2021). Distinguishing barriers to insurance in thai villages. *Journal of Human Resources*.

Kocherlakota, N. R. (1996). Implications of efficient risk sharing without commitment. *The Review of Economic Studies*, 63(4), 595–609.

Laczó, S. (2015). Risk sharing with limited commitment and preferences heterogeneity: Structural estimation and testing. *Journal of the European Economic Association*, 13(2), 265–292.

Lewis, D. J., Melcangi, D., & Pilossoph, L. (2019). *Latent Heterogeneity in the Marginal Propensity to Consume*. Staff Reports 902, Federal Reserve Bank of New York.

Ligon, E. (1998). Risk sharing and information in village economies. *The Review of Economic Studies*, 65(4), 847–864.

Ligon, E. (2020). *Estimating household welfare from disaggregate expenditures*. Working Paper 109, Center for Effective Global Action.

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

Lim, Y. & Townsend, R. M. (1998). General equilibrium models of financial systems: Theory and measurement in village economies. *Review of Economic Dynamics*, 1(1), 59–118.

Mace, B. J. (1991). Full insurance in the presence of aggregate uncertainty. *Journal of Political Economy*, 99(5), 928 – 956.

Mazzocco, M. & Saini, S. (2012). Testing efficient risk sharing with heterogeneous risk preferences. *American Economic Review*, 102(1), 428–68.

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

Murphy, K. M. & Topel, R. H. (1985). Estimation and inference in two-step econometric models. *Journal of Business I&' Economic Statistics*, 3(4), 370–379.

Rogerson, W. P. (1985). Repeated moral hazard. *Econometrica*, 53(1), 69–76.

Rosenzweig, M. R. & Binswanger, H. P. (1993). Wealth, weather risk and the composition and profitability of agricultural investments. *Economic Journal*, 103(416), 56 – 78.

Rosenzweig, M. R. & Stark, O. (1989). Consumption smoothing, migration, and marriage: Evidence from rural india. *Journal of Political Economy*, 97(4), 905 – 926.

Rosenzweig, M. R. & Wolpin, K. I. (1993). Credit market constraints, consumption smoothing, and the accumulation of durable production assets in low-income countries: Investment in bullocks in india. *Journal of Political Economy*, 101(2), 223 – 244.

Schulhofer-Wohl, S. (2011). Heterogeneity and tests of risk sharing. *Journal of Political Economy*, 119(5), 925–958.

Shrinivas, A. & Fafchamps, M. (2018). Testing efficient risk sharing with heterogeneous risk preferences: Comment. *American Economic Review*, 108(10), 3104–13.

Townsend, R. M. (1982). Optimal multiperiod contracts and the gain from enduring relationships under private information. *Journal of Political Economy*, 90(6), 1166–1186.

Townsend, R. M. (1994). Risk and insurance in village india. *Econometrica*, 62(3), 539 – 591.

Udry, C. (1990). Credit Markets in Northern Nigeria: Credit as Insurance in a Rural Economy. *The World Bank Economic Review*, 4(3), 251–269.

Wilson, R. (1968). The theory of syndicates. *Econometrica*, 36, 119 – 132.

World Bank (2014). World development indicators database. data retrieved from https://data.worldbank.org.

Zeldes, S. P. (1989). Consumption and liquidity constraints: An empirical investigation. *Journal of Political Economy*, 97(2), 305–346.

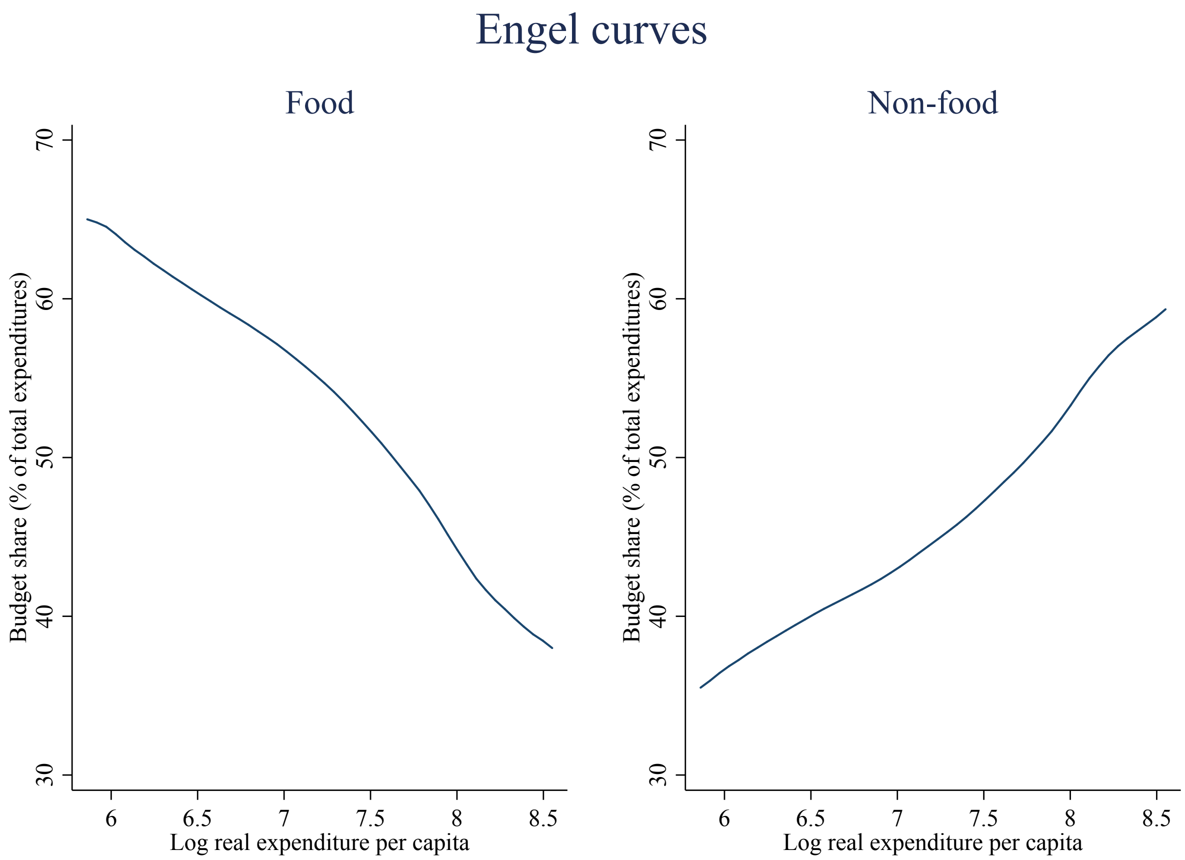

Figure 1: Engel Curves for food and non-food

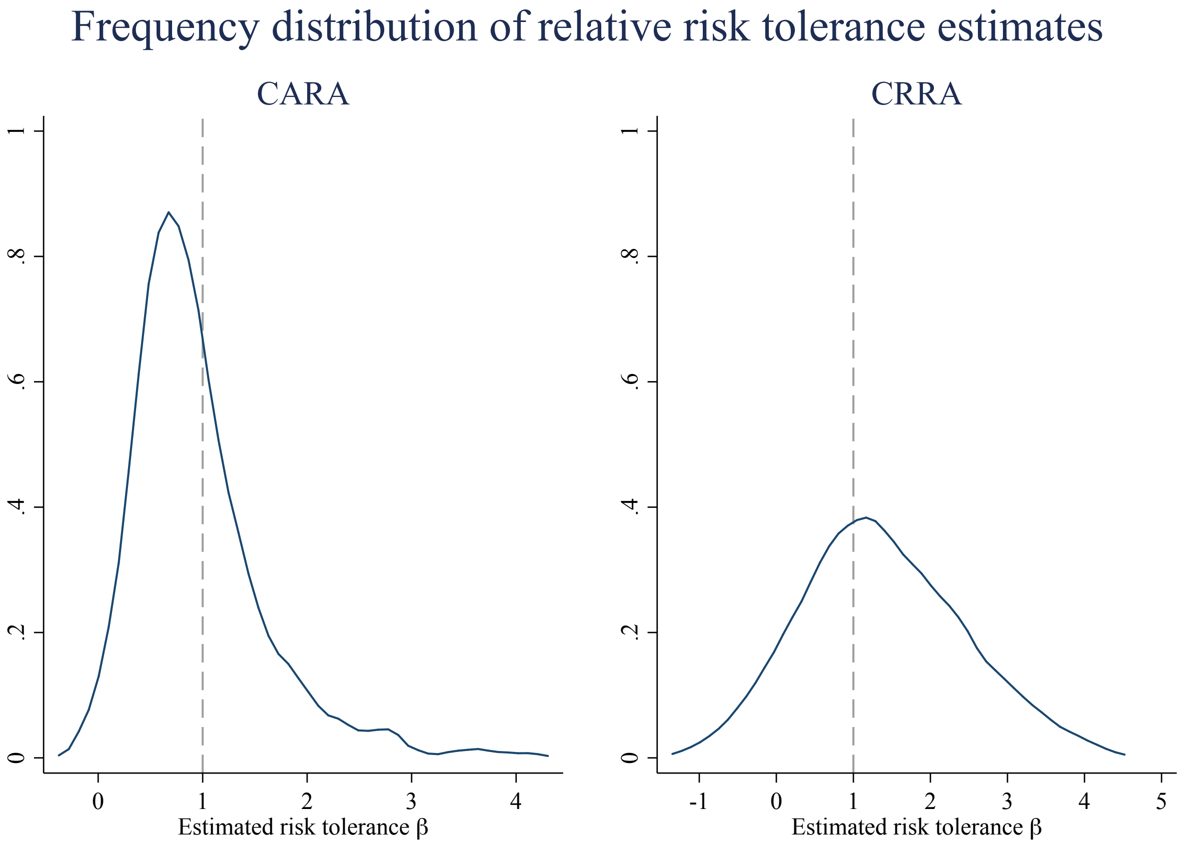

Figure 2: Normalized risk tolerance β

Notes: Each panel of this Figure depicts the frequency distribution of $\hat{\beta}_i$, the estimate of household-specific risk tolerance obtained from regression (10). In each case, the sample mean of $\hat{\beta}_i$ is set to 1 within each village. This means that $\hat{\beta}_i$ measures the risk tolerance of household i relative to other households in the same village.

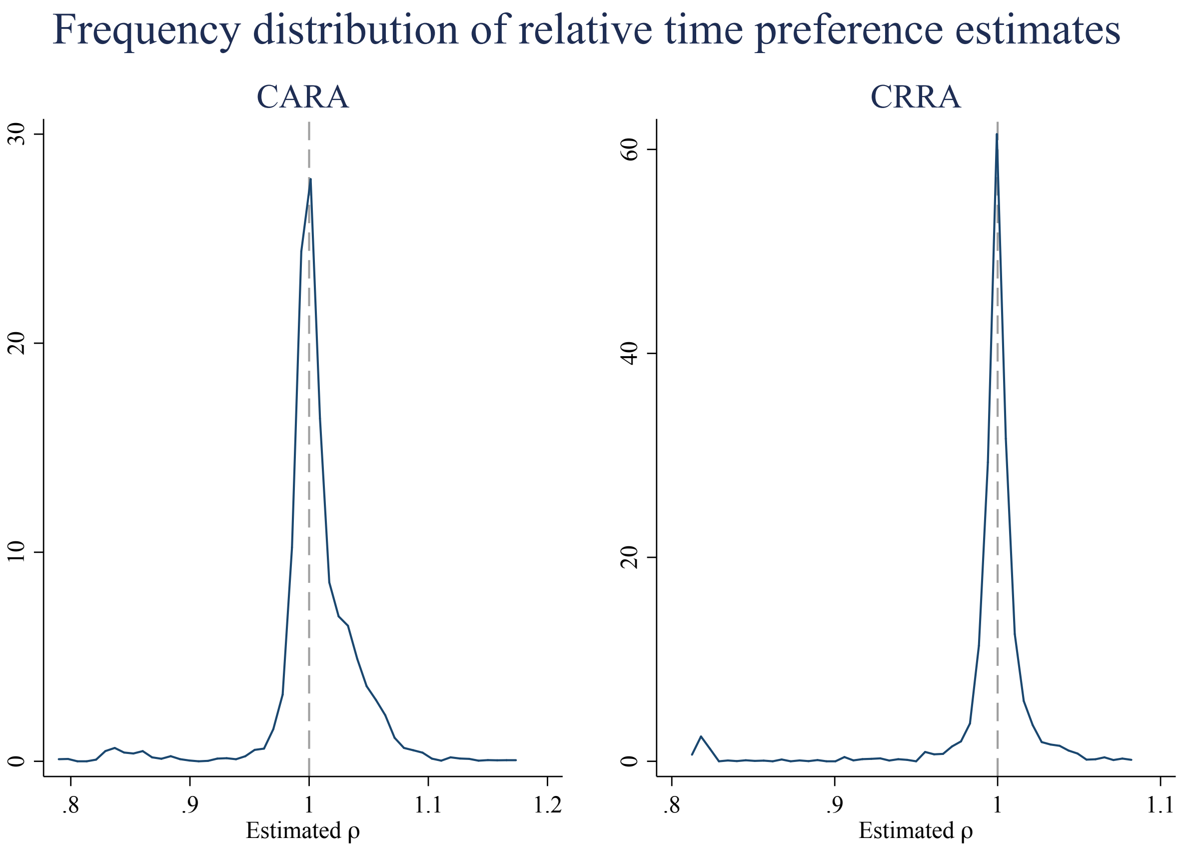

Figure 3: Normalized time preference parameter ρ

Notes: Each panel of this Figure depicts the frequency distribution of $\hat{\rho}_i$, the estimate of household-specific discount factor obtained from regression (10). As for risk tolerance, the sample mean of $\hat{\rho}_i$ is set to 1 within each village – which that $\hat{\rho}_i$ measures the discount factor of household i relative to other households in the same village.

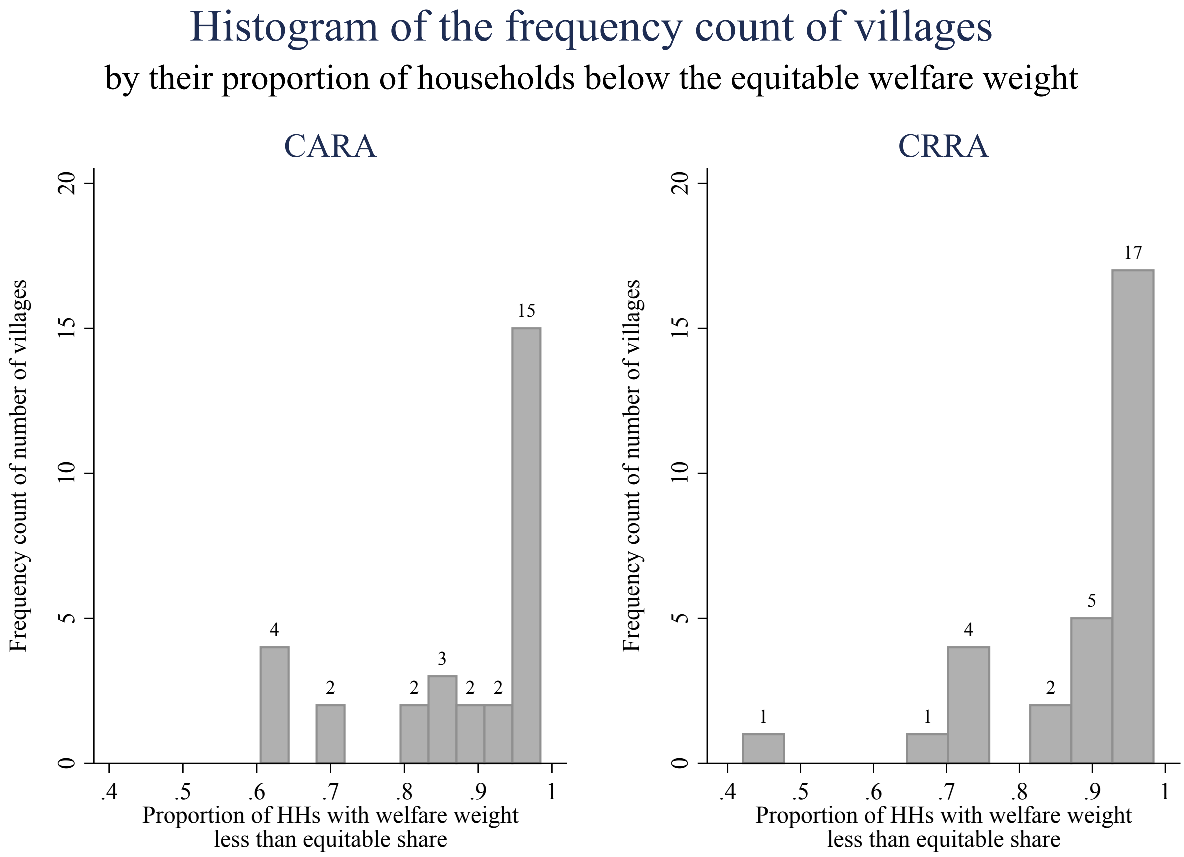

Figure 4: Consumption Inequality

Notes: Each panel of this Figure depicts the histogram of village-average proportion of households with welfare weight $\hat{\eta}_i$ less than the average welfare weight $1/N_v$ for the 30 villages, one from CARA and the other from CRRA. The household-specific welfare weight estimate $\hat{\eta}_i$ is obtained from regression (10). In each case, the sum of $\hat{\eta}_i$ is set to 1 within each village. Based on CARA estimates, the overall mean is 0.88, implying that 88% of HHs have welfare weights less than equitable share. Similarly, the overall mean for CRRA estimates is 0.89.

Table 1: Summary statistics

	Mean	SD
Consumption		
Total expenditure	15030	9754
Food expenditure	8026	3838
Non-food expenditure	6912	6608
Income		
Crop and Livestock income	148	9213
Wages income	1229	1443
Earned income	15477	22028
Assets		
Wealth	84467	101410
Liquid wealth	25958	43571
Cash-in-Hand	30074	45937
Household characterisics		
Household size	4.8	2.3
Adult-eq weight	4.1	1.8
Education of head (in years)	5.1	4.7
Age of head	50	13
Households	1296	
Villages	30	
Observations	6480	

Notes: This table reports descriptive statistics - mean and standard deviation - for household consumption, income, assets and demographic characteristics. Consumption, Income and Asset variables represent annual values, adjusted to 2010 rupees per adult-equivalent. Total expenditures is the sum of food and non-food expenditures, and earned income is the sum of crop and livestock income and wage income. Wealth is sum of net credit position (saving-borrowing+lending) and total assets (liquid assets+capital assets); Liquid wealth is the sum of net credit position (saving-borrowing+lending) and liquid assets (livestock, consumer durables and inventory value of crops, inputs and fuel). Similarly, cash-in-hand is constructed as the sum of liquid wealth and income earned. For all values, top and bottom 1% is trimmed for measurement errors.

Table 2: First differences in levels - CARA model

		Homothetic preferences		Non-homothetic pref	
Expenditures:	Food+Non-food (1)	Food (2)	Non-food (3)	Food (4)	Non-food (5)
Panel A: Homogeneous preferences					
Average village expenditures	0.9236*** (0.0396)	0.5878*** (0.0603)	0.9041*** (0.0609)	0.5926*** (0.0971)	0.7693*** (0.0578)
Liquid wealth	0.0108** (0.0040)	0.0068** (0.0031)	0.0140** (0.0056)	0.0084** (0.0040)	0.0103*** (0.0037)
Earned Income	0.0314*** (0.0097)	0.0289*** (0.0078)	0.0322** (0.0118)	0.0288*** (0.0098)	0.0237*** (0.0077)
Observations	5028	5028	5028	5028	5028
Panel B: Heterogeneous preferences					
Liquid wealth	0.0106*** (0.0037)	0.0054* (0.0030)	0.0137** (0.0053)	0.0070* (0.0036)	0.0095** (0.0036)
Earned Income	0.0292*** (0.0093)	0.0235*** (0.0070)	0.0298** (0.0118)	0.0234** (0.0090)	0.0200** (0.0080)
Observations	5028	5028	5028	5028	5028

Notes: This table reports the results from household panel-pooled estimation of first differences in levels, based on CARA model. All variables are in 2010 rupees per adult equivalent, aggregated over the entire year. The unit of observation is a household-year. Each column presents the results from a separate regression on a different dependent variable. Column and row headings correspond to the dependent and regressor variables, respectively. Panel A reports the test results of full risk pooling under homogeneous preferences and Panel B reports the test results of full risk pooling after accounting for heterogeneity in risk and time preferences. Village expenditures represents the village leave-out-mean. Liquid wealth is the sum of the net credit position (saving-borrowing+lending) and liquid assets (livestock, consumer durables and inventory value of crops, inputs and fuel). Earned income is the sum of income from crop, livestock and wages. The results under the two columns 'Homothetic preferences' assumes linear Engel curves: food and non-food expenditures are divided by the sample average budget share. In the two columns 'Non-homothetic preferences', food and non-food expenditures represent the predicted total expenditures that correspond to a particular food or non-food level of expenditures, based on sample-estimated Engel curves. In Panel B, the dependent variable is transformed to account for heterogeneity in risk and time preferences – see text for details. The top and bottom 1% of all values are trimmed to eliminate outliers and large measurement errors. All standard errors are clustered at the village level. * p<0.10 ** p<0.05 *** p<0.01.

Table 3: First differences in logs - CRRA model

		Homothetic preferences		Non-homothetic pref	
Expenditures:	Food+Non-food (1)	Food (2)	Non-food (3)	Food (4)	Non-food (5)
Panel A: Homogeneous preferences					
Average village expenditures	0.9094*** (0.0241)	0.5780*** (0.0596)	1.2524*** (0.1499)	0.3555*** (0.0495)	0.7959*** (0.0935)
Liquid wealth	0.0263*** (0.0048)	0.0214*** (0.0067)	0.0227** (0.0100)	0.0267*** (0.0057)	0.0166** (0.0064)
Earned Income	0.0182*** (0.0057)	0.0220** (0.0084)	0.0160* (0.0087)	0.0184** (0.0074)	0.0122** (0.0057)
Observations	3185	3172	3181	3185	3185
Panel B: Heterogeneous preferences					
Liquid wealth	-0.0010 (0.0089)	-0.0098 (0.0122)	0.0008 (0.0097)	-0.0092 (0.0127)	-0.0124 (0.0087)
Earned Income	0.0093 (0.0088)	0.0054 (0.0120)	0.0133 (0.0100)	-0.0004 (0.0125)	0.0013 (0.0098)
Observations	3185	3172	3181	3185	3185

Notes: This Table reports the results from a household panel pooled estimation of first differences in logs, based on the CRRA model. All variables are in 2010 rupees per adult equivalent, aggregated over the entire year. The unit of observation is a household-year. Each column presents the results from a separate regression on a different dependent variable. Column and row headings correspond to the dependent and regressor variables, respectively. Panel A reports the test results of full risk pooling under homogeneous preferences and Panel B reports the test results of full risk pooling after accounting for heterogeneity in risk and time preferences. Village expenditures represents the village leave-out-mean. Liquid wealth is the sum of the net credit position (saving-borrowing+lending) and liquid assets (livestock, consumer durables and inventory value of crops, inputs and fuel). Earned income is the sum of income from crop, livestock and wages. The results under the two columns 'Homothetic preferences' assumes linear Engel curves: food and non-food expenditures are divided by the sample average budget share. In the two columns 'Non-homothetic preferences', food and non-food expenditures represent the predicted total expenditures that correspond to a particular food or non-food level of expenditures, based on sample-estimated Engel curves. In Panel B, the dependent variable is transformed to account for heterogeneity in risk and time preferences – see text for details. The top and bottom 1% of all values are trimmed to eliminate outliers and large measurement errors. All standard errors are clustered at the village level. * p<0.10 ** p<0.05 *** p<0.01.

Table 4: Village precautionary savings

		Homothetic pref.		Non-homothetic pref.	
Expenditures:	Food+Non-food (1)	Food (2)	Non-food (3)	Food (4)	Non-food (5)
Panel A: First differences					
Cash-in-hand	0.0309**	0.0187	0.0486**	0.0211	0.0320**
	(0.0146)	(0.0132)	(0.0204)	(0.0150)	(0.0146)
Income earned	0.1189**	0.0913**	0.1366**	0.1118**	0.0885*
	(0.0470)	(0.0422)	(0.0655)	(0.0482)	(0.0469)
R-squarred	0.074	0.047	0.068	0.052	0.058
Observations	120	120	120	120	120
Panel B: Log differences					
Cash-in-hand	0.0406	0.0178	0.0573	0.0189	0.0372
	(0.0257)	(0.0227)	(0.0484)	(0.0244)	(0.0317)
Income earned	0.0948**	0.0876**	0.0915	0.1021**	0.0626
	(0.0440)	(0.0389)	(0.0829)	(0.0419)	(0.0543)
R-squarred	0.059	0.048	0.023	0.055	0.023
Observations	116	116	116	116	116

Notes: This table reports the test of precautionary saving at the village level. Each column presents the results from a separate regression on a different dependent variable. Column and row headings correspond to the dependent and regressor variables, respectively. The unit of observation is a village-year and all variables represent village averages. All variables are in 2010 rupees per adult equivalent, aggregated over the entire year. Panel A reports the test results for first differences in levels and Panel B reports the results for first difference in logs. Liquid wealth is the sum of the net credit position (saving-borrowing+lending) and liquid assets (livestock, consumer durables and inventory value of crops, inputs and fuel). Earned income is the sum of income from crop, livestock and wages. The results under the two columns 'Homothetic preferences' assumes linear Engel curves: food and non-food expenditures are divided by the sample average budget share. In the two columns 'Non-homothetic preferences', food and non-food expenditures represent the predicted total expenditures corresponding to a particular food or non-food level of expenditures, based on sample-estimated Engel curves. Standard errors are reported in parenthesis. * p<0.10 ** p<0.05 *** p<0.01.

Table 5: Risk pooling within sub-castes

	CARA Model : First difference in levels					CRRA model : First difference in log units				
		Homothetic preferences		Non-homothetic pref			Homothetic preferences		Non-homothetic pref	
Expenditures:	Food+Non-food (1)	Food (2)	Non-food (3)	Food (4)	Non-food (5)	Total exp (6)	Food (7)	Non-food (8)	Food (9)	Non-food (10)
Panel A: Homogeneous preferences										
Caste Expenditure	0.7129***	0.4305***	0.8014***	0.4289***	0.5894***	0.7845***	0.4845***	1.0723***	0.5533***	1.1103***
	(0.0351)	(0.0450)	(0.0752)	(0.0644)	(0.0547)	(0.0298)	(0.0494)	(0.0985)	(0.0418)	(0.1137)
Liquid wealth	0.0126***	0.0076**	0.0177***	0.0086**	0.0118***	0.0259***	0.0186***	0.0237**	0.0256***	0.0247***
	(0.0039)	(0.0031)	(0.0067)	(0.0039)	(0.0041)	(0.0051)	(0.0066)	(0.0091)	(0.0063)	(0.0075)
Income earned	0.0340***	0.0315***	0.0400**	0.0314***	0.0259***	0.0158***	0.0198**	0.0134	0.0207***	0.0101
	(0.0115)	(0.0086)	(0.0161)	(0.0107)	(0.0095)	(0.0054)	(0.0082)	(0.0099)	(0.0064)	(0.0098)
Observations	4336	4336	4336	4336	4336	2730	2721	2726	2734	2466
Panel B: Heterogeneous preferences										
Liquid wealth	0.0104***	0.0052*	0.0156**	0.0061*	0.0095**	0.0110	-0.0003	0.0133	0.0028	0.0008
	(0.0036)	(0.0027)	(0.0065)	(0.0034)	(0.0039)	(0.0069)	(0.0092)	(0.0097)	(0.0094)	(0.0069)
Income earned	0.0327***	0.0262***	0.0400***	0.0262***	0.0230**	0.0144	0.0111	0.0179	0.0059	0.0071
	(0.0103)	(0.0076)	(0.0151)	(0.0096)	(0.0090)	(0.0098)	(0.0073)	(0.0113)	(0.0076)	(0.0086)
Observations	4336	4336	4336	4336	4336	2730	2721	2726	2721	2726

This table reports the results from household panel-pooled estimation of first differences in levels and logs, based on the CARA and CRRA models, respectively. The left-hand panel is similar to Table 2, but estimated at the Jati or sub-caste level. The right-hand panel is similar to Table 3, but estimated at the Jati or sub-caste level. All variables are in 2010 rupees per adult equivalent, aggregated over the entire year. The unit of observation is a household-year. Each column presents the results from a separate regression on a different dependent variable. Column and row headings correspond to the dependent and regressor variables, respectively. Panel A reports the test results of full risk pooling under homogeneous preferences and Panel B reports the test results of full risk pooling after accounting for heterogeneity in risk and time preferences. Village expenditures represents the village leave-out-mean. Liquid wealth is the sum of the net credit position (saving-borrowing+lending) and liquid assets (livestock, consumer durables and inventory value of crops, inputs and fuel). Earned income is the sum of income from crop, livestock and wages. The results under the two columns 'Homothetic preferences' assumes linear Engel curves: food and non-food expenditures are divided by the sample average budget share. In the two columns 'Non-homothetic preferences', food and non-food expenditures represent the predicted total expenditures that correspond to a particular food or non-food level of expenditures, based on sample-estimated Engel curves. In Panel B, the dependent variable is transformed to account for heterogeneity in risk and time preferences – see text for details. The top and bottom 1% of all values are trimmed to eliminate outliers and large measurement errors. All standard errors are clustered at the village level. * p<0.10 ** p<0.05 *** p<0.01.

Table 6: Caste-level precautionary savings

		Homothetic pref		Non-homothetic pref.	
Expenditures:	Food+Non-food (1)	Food (2)	Non-food (3)	Food (4)	Non-food (5)
Panel A: First differences					
Cash-in-hand	0.0228***	0.0145***	0.0378***	0.0154***	0.0261***
	(0.0062)	(0.0053)	(0.0088)	(0.0059)	(0.0061)
Income earned	0.0854***	0.0705***	0.0873***	0.0699***	0.0591***
	(0.0174)	(0.0149)	(0.0249)	(0.0166)	(0.0173)
R-squarred	0.053	0.043	0.044	0.035	0.042
Observations	659	659	659	659	659
Panel B: Log differences					
Cash-in-hand	0.0263**	0.0326***	0.0269	0.0349***	0.0162
	(0.0120)	(0.0108)	(0.0210)	(0.0105)	(0.0126)
Income earned	0.0361**	0.0268**	0.0409	0.0211*	0.0276*
	(0.0147)	(0.0132)	(0.0256)	(0.0128)	(0.0154)
R-squarred	0.021	0.025	0.008	0.026	0.009
Observations	560	560	560	560	560

Notes: This table reports the test of precautionary saving at the Jati or sub-caste level. Each column presents the results from a separate regression on a different dependent variable. Column and row headings correspond to the dependent and regressor variables, respectively. The unit of observation is a sub-caste-year and all variables represent sub-caste averages. All variables are in 2010 rupees per adult equivalent, aggregated over the entire year. Panel A reports the test results for first differences in levels and Panel B reports the results for first difference in logs. Liquid wealth is the sum of the net credit position (saving-borrowing+lending) and liquid assets (livestock, consumer durables and inventory value of crops, inputs and fuel). Earned income is the sum of income from crop, livestock and wages. The results under the two columns 'Homothetic preferences' assumes linear Engel curves: food and non-food expenditures are divided by the sample average budget share. In the two columns 'Non-homothetic preferences', food and non-food expenditures represent the predicted total expenditures corresponding to a particular food or non-food level of expenditures, based on sample-estimated Engel curves. Standard errors are reported in parenthesis. * p<0.10 ** p<0.05 *** p<0.01. * p<0.10 ** p<0.05 *** p<0.01.

Appendix A: Testing strategy with CRRA preferences

Building on the work of Mace (1991) and Townsend (1994), our testing strategy can easily be extended to the case where individuals have CRRA preferences of the form $U_i(c) = \frac{1}{1-\gamma_i} c^{1-\gamma_i}$ where parameter γ_i is the coefficient of relative risk aversion of individual i. Under CRRA, the FOC for perfect risk sharing by the social planner simplifies to:

$$\eta_i \rho_i^t c^{-\gamma_i} = \lambda_{ts}$$

Taking logs and rearranging, we get:

$$log\, c_{its} = \frac{log\, \eta_i}{\gamma_i} + \frac{log\, \rho_i}{\gamma_i} t - \frac{1}{\gamma_i} log \lambda_{ts} \tag{13}$$

Averaging over all N individuals in the village and solving for $log\lambda_{ts}$ yields an expression for average village log consumption $\overline{log\, c}_{ts} \equiv \frac{1}{N}\sum_{i=1}^{N} log\, c_{its}$, which we use to replace the common Lagrange multiplier in equation (13). We then obtain:

$$log\, c_{its} = \frac{1}{\gamma_i}\left[log\, \eta_i - \frac{\frac{1}{N}\sum_{j=1}^{N}\frac{log\eta_j}{\gamma_j}}{\frac{1}{N}\sum_{j=1}^{N}\frac{1}{\gamma_j}}\right] + \frac{1}{\gamma_i}\left[log\rho_i - \frac{\frac{1}{N}\sum_{j=1}^{N}\frac{log\rho_j}{\gamma_j}}{\frac{1}{N}\sum_{j=1}^{N}\frac{1}{\gamma_j}}\right] t + \frac{1/\gamma_i}{\frac{1}{N}\sum_{j=1}^{N}\frac{1}{\gamma_j}} \overline{log\, c}_{ts} \tag{14}$$

Under homogeneous risk preferences, the regression used to test efficient risk sharing becomes:

$$\Delta log\, c_{it} = \beta_1 + \beta_2 \Delta \overline{log\, c}_t + \beta_3 \Delta log y_{it} + \beta_4 \Delta log w_{it} + \epsilon_{it} \tag{15}$$

where the two exclusion restrictions in levels present in equation (8) have been suitably replaced by their equivalent in logs.

Under heterogeneous preferences, we similarly start by normalizing risk preferences relative to their mean by imposing that $\frac{1}{N}\sum\limits_{i=1}^{N}\frac{1}{\gamma_i} = 1$. With this normalization, we obtain:

$$log\, c_{it} = \frac{1}{\gamma_i}\left[log\, \eta_i - \frac{1}{N}\sum_{j=1}^{N}\frac{log\eta_j}{\gamma_j}\right] + \frac{1}{\gamma_i}\left[log\rho_i - \frac{1}{N}\sum_{j=1}^{N}\frac{log\rho_j}{\gamma_j}\right] t + \frac{1}{\gamma_i}\overline{log\, c}_t$$

To estimate this model, we first need to obtain estimates of all individual γ_i and ρ_i by running a model of the form:

$$log\, c_{it} = \alpha_i + \theta_i t + \beta_i \overline{log\, c_t} + \epsilon_{it} \tag{16}$$

using, as before, monthly consumption data on household i. We then recover structural parameters using the following equalities:

$$\beta_i = \frac{1}{\gamma_i} \tag{16A}$$

$$\alpha_i = \frac{1}{\gamma_i}\left[\log \eta_i - \frac{1}{N}\sum_{j=1}^{N}\frac{\log \eta_j}{\gamma_j}\right] \tag{16B}$$

$$\theta_i = \frac{1}{\gamma_i}\left[\log \rho_i - \frac{1}{N}\sum_{j=1}^{N}\frac{\log \rho_j}{\gamma_j}\right] \tag{16C}$$

In this case, the estimated γ_i can be interpreted as capturing the extent to which the coefficient of relative risk aversion of individual i differs from the average degree of relative risk aversion in the sample. The formulas for recovering welfare weights and time preference parameters are unchanged.

It follows that, apart from a slight difference in the normalization of risk preferences, the estimation of the CRRA and CARA model household by household is very similar, except that, in the CARA case, individual and village consumption are expressed in levels while they appear in logs in the CRRA case. The estimated regression model (10) for the CRRA case is thus:

$$log\, c_{it} - \hat{\beta}_i \overline{log\, c_t} - \hat{\theta}_i t = \xi log\, y_{it} + \zeta log\, w_{it} + \epsilon_{it}$$

or, expressed in first difference:

$$\Delta\, log\, c_{it} - \hat{\beta}_i \Delta\, \overline{log\, c_t} - \hat{\theta}_i = \xi \Delta\, log\, y_{it} + \zeta \Delta\, log\, w_{it} + \epsilon_{it}$$

Appendix B: Accounting for Measurement Error

Because we do not observe all the households in a village, but only a sample, there is a measurement error in the Townsend test that causes a downward bias in β_2. Let the true data generating process be:

$$\begin{aligned} c_i &= \beta_0 + \beta_2 \overline{x}_i + u_i \\ x_i &= \overline{x}_i + e_i \end{aligned}$$

where $\overline{x}_i$ is the true village mean for individual i, x_i is a sample mean of $\overline{x}_i$ based on a sample of size N, and e_i is the measurement error. Under the null of perfect risk sharing, the estimated model is:

$$c_i = \beta_0 + \beta_2^* x_i + u_i$$

and the magnitude of the bias is given by:

$$E[\beta_2^*] = \beta_2 (1 - \frac{\sigma_e^2}{\sigma_{\overline{x}}^2})$$

Since the standard error of a sample mean is $\sigma_e = \frac{\sigma_{\overline{x}}}{\sqrt{N}}$, the downward bias is approximately:

$$\begin{aligned} E[\beta_2^*] &= \beta_2 (1 - \frac{\sigma_{\overline{x}}^2 / N}{\sigma_{\overline{x}}^2}) \\ &= \beta_2 (1 - \frac{1}{N}) \\ &\simeq \beta_2 0.976 \end{aligned}$$

when using the full sample of 1296 households divided into 30 villages, i.e., N =42.16 on average across villages. We can use this result to perform an approximate correction of the $\hat{\beta}_2$ coefficients estimated for total consumption, i.e., by dividing $\hat{\beta}_2$ by 0.976. We can check whether, as the result of this correction, the revised $\hat{\beta}_2$ is close enough to 1 to fail to reject full risk pooling. A similar calculation for the caste regressions yields a correction factor of 0.946. In both cases, we see that the correction for sampling error is not large enough to qualitatively change our reported findings.

Appendix C: Additional Figures

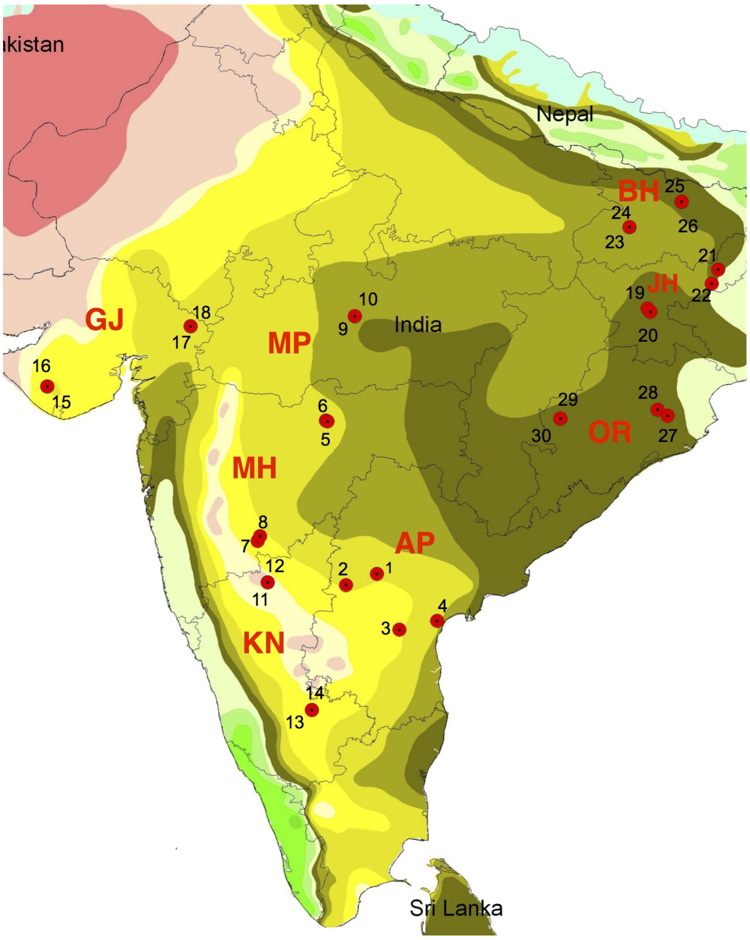

Figure A1: Location of ICRISAT VDSA villages - 30 villages across 8 states

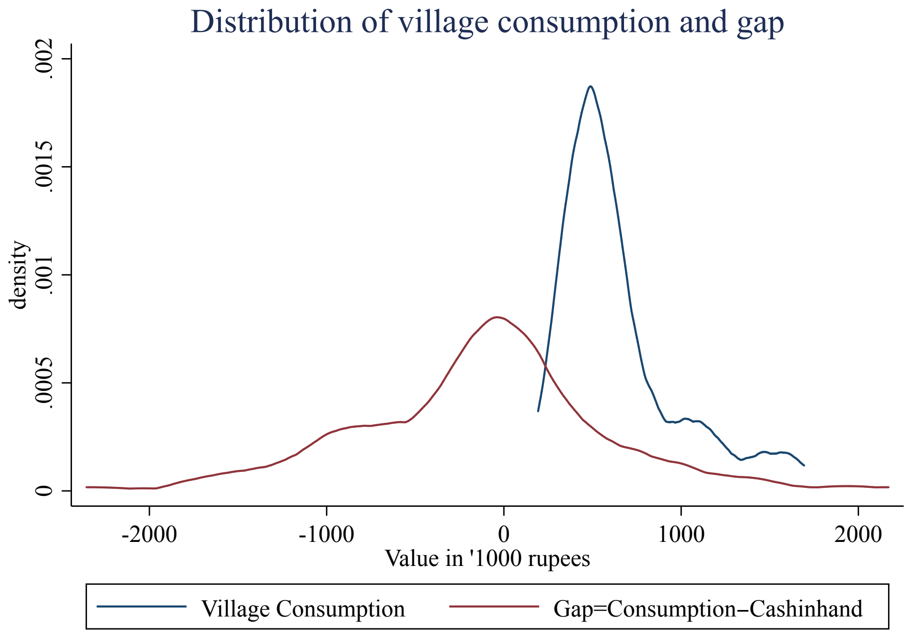

Figure A2: Distribution of Village Consumption and Unexplained Gap