Intertemporal Risk Pooling in Village Economies*

September 2025

Abstract

We propose an improved theoretically-grounded method to test for efficient risk pooling that allows for intertemporal smoothing, non-homothetic consumption, and heterogeneous risk and time preferences. Applying this method to recent panel data from Indian villages generates important new insights while confirming some earlier findings. Year-to-year smoothing of consumption takes place much more at the village level than at the individual level and occurs primarily through financial assets. While there is proportionally more smoothing of food than non-food consumption, accounting for differences in income elasticities between the two statistically eliminates this difference, indicating that risk pooling does not distort consumption choices in our study area. Finally, we find that consumption smoothing is affected jointly by income and liquid assets, and that there is no excess sensitivity to earned income.

Keywords: Risk pooling, precautionary saving, non-homotheticity, heterogeneous risk and time preferences, India

*We thank Pierre Dubois, Cynthia Kinnan, and Pau Milan for their comments, as well as the participants to the WEFIDEV webinar of November 2021 and to the IZA Workshop on Gender and Family Economics of June 2022. We are grateful to ICRISAT for sharing the data with us from their website http://vdsa.icrisat.ac.in. The only input into this research is the time of the two researchers involved. No outside funding was secured for this project. We have no conflict of interest to report.

1 Introduction

People in village economies face 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 and Gruber, 2002; De Weerdt and 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 (Zeldes, 1989a; Deaton, 1991). Risk can also be pooled *ex post* through various informal or formal agreements (Udry, 1990; Fafchamps and Lund, 2003; Dercon et al., 2006). Risk pooling is best at addressing idiosyncratic shortfalls in income that only affect some households. 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). It follows that effective protection of village economies against income shocks requires a combination of risk sharing and precautionary saving.

In spite of this, the literature on consumption smoothing has tended to focus on either of these two mechanisms in isolation. Early studies of consumption smoothing across individuals have focused on risk pooling within periods (e.g., Mace, 1991; Altonji et al., 1992; Townsend, 1994) 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, 2022). Similarly, studies of intertemporal risk coping via precautionary savings have typically assumed away contemporaneous risk pooling across households (e.g., Zeldes, 1989a; Deaton, 1991; Rosenzweig and Wolpin, 1993; Lim and Townsend, 1998; Fafchamps et al., 1998; Kazianga and Udry, 2006).

In this paper, we propose a novel approach to testing for efficient risk pooling that combines both strands of the literature into a single theoretically-based framework. Its purpose is to test whether village economies optimally pool risk without unrealistically assuming away wealth accumulation, as tests of optimal risk pooling have done to date. Perfect risk pooling at the village level is calculated by solving a social planner optimization that maximizes the weighted sum of discounted expected utilities of all households subject to an intertemporal budget constraint. In this framework, the decision of the social planner can be divided into two steps: in each period, the social planner first optimally chooses aggregate village consumption as in a precautionary saving model (Zeldes, 1989a); in a second step, this aggregate consumption is optimally divided among households (Townsend, 1994).[1] This framework allows for individual asset accumulation but, in an efficient equilibrium, all assets are de facto held in trust for the entire risk pooling group and all village savings and income are pooled to smooth consumption over time and within periods. In such equilibrium, the optimal division of wealth across villagers

[1] Past empirical tests of risk pooling are nested in our model and, as such, are unbiased – although they may be underpowered due to their reliance on income alone in their exclusion restriction, ignoring the possible role of liquid assets. On the other hand, 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. Hence precautionary saving can only be meaningfully studied at the aggregate level of the risk pooling group.

only depends on whether returns to individual assets are decreasing, increasing, or constant.[2] This is because individually held assets can be used to smooth the consumption of others. This framework bears some similarity with how lending by saving and credit associations to their members can help non-members and enhance risk pooling (Attanasio et al., 2023).

As a second main innovation, we extend the above model to multiple consumption categories and use it to develop a new risk sharing test that allows for non-homothetic consumption preferences. As noted in Mace (1991), if households have homothetic consumption preferences, perfect risk pooling implies that all consumption categories co-vary equally to shocks: as the total expenditures of the household increase or contract, all consumption categories should rise and fall in the same proportion. However, when preferences are non-homothetic, the consumption of goods with an income elasticity larger than one should co-move more than proportionally to a rise in total expenditures, while the opposite holds for goods with a low income elasticity. We propose a new method of using partial consumption data to test for efficient risk pooling that accounts for non-homothetic preferences. Our method involves estimating Engel curves on cross-section data and predicting by how much expenditures of different expenditure categories should vary with total expenditures to remain on the household's indifference curve. This approach allows us to test whether certain components of consumption are better insured than others, as could arise, for instance, in the presence of paternalism or imperfect altruism.

In addition, we allow for heterogeneity in risk and time preferences among households, as has successfully been done in the recent literature (e.g., Schulhofer-Wohl, 2011; Mazzocco and Saini, 2012; Chiappori et al., 2014). This approach 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. We follow the method proposed by Chiappori et al. (2014), and estimate relative risk preferences under the assumption of perfect risk pooling, and then use the preference estimates to correct the test for full risk pooling. Finally, we replicate the analysis assuming aggregation of risk either at the village-level or the sub-caste level.

We use the new wave of ICRISAT panel data that cover 30 villages in India between 2010 to 2015 to test for risk pooling in Indian villages. We find evidence of less than perfect risk pooling at the village level, for total expenditures. These results are qualitatively similar to earlier studies that test for risk pooling using the old ICRISAT data that cover six villages between 1975 to 1985 (Townsend, 1994). But this shortfall is quite limited in the sense that household consumption expenditures move nearly one-for-one with the village average. Even though year-to-year variation in household-level cash-in-hand has 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

[2]If returns are convex, all precautionary assets should be held by a single actor (e.g., an insurance corporation or a large money lender). If they are concave, wealth should be equally distributed so as to maximize aggregate returns. If the saving technology has linear returns (e.g., financial savings), the distribution of assets does not matter.

risk.[3] Our results also confirm previous findings indicating that there is substantial and significant heterogeneity in estimated risk preferences across households – and we show that this conclusion extends to time preferences. As previous authors have argued theoretically, we find that 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.

Our first novel finding relates to consumption smoothing across years. We find that individual household consumption varies significantly but weakly with their own cash-in-hand, whereas village consumption varies more strongly to village cash-in-hand. These results suggest that year-to-year smoothing of consumption takes place much more at the village level than at the individual level. Furthermore, we find that villages optimally draw on the most liquid assets - financial assets and jewelry (i.e., gold and silver) - to smooth aggregate village income shocks. Taken together, these findings suggest that our study villages manage to achieve a considerable smoothing of consumption in response to idiosyncratic within-year income shocks and that they use liquid wealth – primarily financial assets and jewelry – to smooth village consumption across years.

Our second novel finding is to show that food and non-food expenditures of individual households vary with average village consumption in a way consistent with a within-year utility-maximizing allocation of total expenditures, once we account for non-linear Engel curves. Under homothetic preferences, we find that food expenditures vary less strongly with village average expenditures and with household cash-in-hand, which may suggest that food consumption may be better insured against aggregate shocks relative to non-food expenditures.[4] However, assuming homothetic preferences may over-estimate the excess sensitivity of non-food expenditures to village shocks relative to food, as non-foods have a higher income elasticity. After we correct for non-homothetic preferences, while we continue to observe that non-foods vary more strongly with aggregate village shocks than food expenditures, the difference is not statistically significant.

Finally, we largely replicate these findings at the sub-caste (*Jati*) level, instead of the village level. We find that, contrary to the papers using earlier Indian data (e.g., Townsend, 1994; Mazzocco and Saini, 2012; Shrinivas and Fafchamps, 2018), risk pooling within sub-castes in each village is less strong – i.e., less responsive to aggregate consumption and more responsive to

[3] 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. 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.

[4] In the context of risk pooling, this is consistent with altruistic paternalism in social welfare interventions in high income countries that favor the consumption of specific goods (e.g., food stamps; subsidized health care and education; affordable housing initiatives) (Currie and Gahvari, 2008; Cunha, 2014; Hastings and Shapiro, 2018). A similar pattern may be present in a low-income informal village setting if risk pooling depends on people's goodwill and is affected by social norms about redistribution (Fafchamps, 1992). 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.

individual income and cash-in-hand – than pooling across all households in the village. Similar conclusions are reached for precautionary saving: within-jati aggregate sensitivity to cash-in-hand is smaller than what we found for within-village consumption. Taken together, these results suggest that, in our study sample, risk pooling takes place more at the village than jati level.

This paper makes several contributions to the literature. First, we estimate a new risk pooling test that accounts for precautionary savings. Previous studies of consumption smoothing have either focused solely on inter-household risk pooling (Mace, 1991; Townsend, 1994; Ligon, 1998; Kinnan, 2022) or inter-temporal precautionary savings (Zeldes, 1989a; Deaton, 1991). We show how the two empirical approaches can be combined within a single framework, and we go well beyond existing studies in the scope of our findings. Our results suggest that consumption smoothing is achieved from within-period risk pooling and year-to-year precautionary savings. The findings indicate that household consumption is largely insulated from individual and collective income shocks.

Second, we integrate non-homothetic Engel curves into our testing strategy. There has been a recent revival of interest in Engel curves (e.g., Dunbar et al., 2013; Atkin et al., 2020; Ligon, 2020; Almås et al., 2018; Escanciano et al., 2021). Drawing inspiration from this literature, we propose a new way of using partial consumption data to test for efficient risk pooling when income elasticities are non-unitary. After correcting for non-homotheticity, we do not find any significant evidence that food consumption is over-insured in the sense that it co-moves less (and non-food consumption co-moves more) with village consumption than is predicted by Engel curves. This suggests that households are able to spend their consumption budget optimally.

Third, this paper adds to the new strand of literature on risk sharing with heterogeneous preferences (e.g., Schulhofer-Wohl, 2011; Mazzocco and Saini, 2012; Chiappori et al., 2014; Dubois, 2001). 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 building 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.[5] Our approach, however, differs from Chiappori et al. (2014) in important ways: it generalizes the approach by simultaneously estimating relative welfare weights and time preferences between households under perfect risk pooling; and it is easier to implement since in relies on a simple linear regression.[6]

[5] Since we test efficient risk pooling, it is natural that preference heterogeneity be estimated under that assumption as well, so as to ensure internal consistency of the test. This being said, Chiappori et al. (2014) show that the bias caused by a violation of this assumption is minimal, and the estimates predominantly identify the true preference parameters. We revisit this point in Section 3.3.

[6] Our approach employs a simple linear regression to estimate the correlation between household consumption

2 Risk pooling with Assets: A conceptual framework

2.1 Adding liquid assets to the standard risk pooling model

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.[7] 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 w_{it} is the stock of liquid assets of individual i saved in the previous period and available at the beginning of period t, and r_s is the return to those assets, which is allowed to vary with the state of the world s. Each 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:

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

with that of aggregate village consumption for each household. Our method differs from previous studies in several ways. Chiappori et al. (2014) and Mazzocco and Saini (2012) consider household-pairs and estimate correlation of consumption between each pair of households in the village. Chiappori et al. (2014) use moment conditions generated by consumption of each pair of distinct households in a village to impute preference parameters of each household by GMM. Mazzocco and Saini (2012) test for efficient risk sharing for household pairs by examining whether a household's consumption is monotonically increasing with the sum of consumption for the pair. Their non-parametric test uses a risk-sharing function that allows for a general class of utility functions with heterogeneous risk preferences. On the other hand, Schulhofer-Wohl (2011) treats risk preferences as nuisance parameters that must be eliminated from the full risk sharing equation. To this effect, the author uses quasi-fixed effects that controls for household specific trends and household specific effects of aggregate shocks, thereby removing any heterogeneity in preferences.

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

where η_i is a particular set of (time-invariant) welfare weights with $\sum\limits_{i=1}^{N} \eta_i = 1$. 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 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 $y_{ts} = \sum\limits_{i}^{N} y_{its}$ 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 $w_{ts} \equiv \sum\limits_{i}^{N} w_{its}$ 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 solution 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:

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

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 weighed by the welfare weights.[8] 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 Bellman equation:

$$V_t(x_{ts}) = max_{w_{t+1,s}} W_t(x_{ts} - w_{t+1,s}) + R(t) E V_{t+1}((1 + r_{s'}) w_{t+1,s'} + y_{ts'}) \tag{4}$$

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 (e.g., Stokey and Lucas, 1989). 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, 1989a; Deaton, 1991). It is well known that $c_{ts} = C_t(x_{ts})$ is a concave function of x_{ts}.

2.2 Interpretation

In order to correctly interpret results from the above risk pooling model, it is essential to understand the logic and purpose of the testing methodology. First, it does not constitute a *causal* analysis of any kind. The theory makes predictions regarding specific correlations patterns that should or should not be present in the data and we test for the presence of these patterns. These tests rest heavily on a simple accounting identity: consumption has to be financed somehow. They do not imply – or require – any causal process. Identification of the risk pooling test nonetheless requires that that cash-in-hand x_{its} be outside the reach of the social planner. If it was not, the social planner could achieve the optimal consumption allocation vector $\{c_{its}\}$ simply by setting the cash-in-hand vector $\{x_{its}\}$ equal to $\{c_{its}\}$. As a result, we would have $c_{its} = x_{its}$ for all i, t, s, implying that the exclusion restriction would not hold and our testing strategy would fail. For identification, cash-in-hand must be at least partly pre-determined from the point of view of the consumption redistribution process. To satisfy this requirement, we follow the rest of the literature by restricting y_{its} to earned sources of income from farming and casual agricultural work, thereby omitting transfers from friends

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

and relatives that could be a response to a negative income shock. We also construct the liquid wealth component of cash-in-hand x_{its} using only liquid assets at the end of the previous period $t-1$.

Second, we restrict our attention to consumption smoothing after *income* smoothing has taken place. It does not matter for the validity of the tests whether or not individual households have already reallocated resources in response to an income shortfall – e.g., by increasing their supply of labor to the market (e.g., Kochar, 1999). The income shocks we consider here are those that the household did not absorb to keep disposable income unchanged. As long as the household is unable to completely smooth income on its own, the need for consumption smoothing remains and it must be achieved by other means, which are the focus of this paper.

Third, we have followed other papers in using the word 'risk' throughout. It is important to realize, however, that the tests we conduct do not require income shocks to be unanticipated or unpredictable. The need for consumption smoothing applies even for perfectly predictable income shortfalls, such as seasonal income variation or the birth of a child. While other papers have attempted to empirically distinguish between anticipated and unanticipated variation income (e.g., Paxson, 1992, 1993; Kochar, 1999; Fafchamps and Lund, 2003), no such attempt is made here because our focus is elsewhere.

Finally, and perhaps most importantly, the methodology followed here does *not*, by itself, identify the mechanisms by which pooling/sharing is achieved. In particular, evidence of risk pooling at the village level does not imply the existence of a *deliberate* mechanism, whether formal or informal, for sharing risk among villagers. Indeed it has long been recognized that individual precautionary saving can closely approximate efficient risk pooling if individuals have sufficient liquid wealth: household with a shortfall cover their income deficit by dis-saving to smooth consumption while households with an income surplus save it for a rainy day (e.g., Deaton, 1992; Pan, 2009).[9] This being said, in poor village economies such as those studied here, the historical lack of liquid savings instruments has made it difficult if not impossible for households to accumulate wealth for consumption smoothing purposes. For this reason, it is commonly believed that the explicit sharing of monetary and food resources among villagers plays an important role in how risk pooling takes place (e.g., Rosenzweig and Stark, 1989; Ligon et al., 2002; Pan, 2009; Kinnan and Townsend, 2012; Samphantharak and Townsend, 2018). An important contribution of this paper is to revisit the role of precautionary savings with more contemporary data from rural India.

[9] Sargent (1987), Chapter 3, provides a simple illustration of this fact. Imagine an isolated island in which people consume a non-storable food, e.g., ripening bananas from their orchard. Some days, more bananas ripen, other days less. The banana harvest is also affected by common shocks, such as hurricanes. Each household is endowed with a stock of liquid wealth in the form of currency. With this currency, they purchase bananas when their daily harvest is insufficient and sell bananas when they have more than they need. The price of bananas equilibrates the market such that, when there are few bananas in aggregate, the price is higher, and vice versa if there are bananas in abundance. In this economy, households with excess bananas de facto transfer them to households with a banana deficit by trading bananas against currency. As long as their stock of currency is sufficient, households can perfectly smooth their consumption against idiosyncratic shocks and absorb common shocks by reducing consumption when the price of bananas is high. In this example, islanders pool risk while thinking they are self-insuring. Deaton (1992) provides an illustration of this model using data from Cote d'Ivoire.

The validity of the tests nonetheless requires a number of conditions to be satisfied: the stock of assets w_{it} is measured before, not after, consumption during time interval t, so as to omit asset sales and loans that help the household smooth consumption at t (e.g., Zeldes, 1989a; Dubois, 2001); our definition of income does not include transfers and remittances received from – or given to – other households during period t since these often serve the purpose of smoothing consumption; we use first differences to control for household time-invariant characteristics that are correlated with its average consumption (e.g., Ravallion and Chaudhuri, 1997); and we observe a large enough proportion of households in a village to construct a sufficiently accurate measure of average village consumption (e.g., Suri, 2011). In the next sub-section, we discuss in more detail the implications of the risk pooling model when individuals differ in their risk and time preferences.

2.3 Accounting for heterogeneous time and risk preferences

Under heterogeneous time preferences, 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 planner's discount factor $R(t)$ puts more weight on impatient individuals. As time passes, their weight in the average $\sum\limits_{i=1}^{N} \eta_i \rho_i^t$ 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 may play in the presence of preference heterogeneity.

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.[10] 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:

$$c_{its} = \frac{log\, \eta_i}{\gamma_i} + \frac{log\, \rho_i}{\gamma_i} t - \frac{1}{\gamma_i} log \lambda_{ts} \tag{5}$$

Averaging over all N individuals in the village and solving for $log\lambda_{ts}$ yields an expression for average village consumption $\bar{c}_{ts} \equiv \frac{1}{N}\sum_{i=1}^{N} c_{its}$, which we use to replace the common Lagrange

[10] Assuming constant relative risk aversion (CRRA) yields a similar result, except that estimating equations are expressed in logs rather than levels. See Appendix C for a formal derivation

multiplier in equation (5). We obtain:

$$c_{its} = \frac{1}{\gamma_i}\left[log\,\eta_i - \frac{\frac{1}{N}\sum_{j=1}^{N}\frac{log\eta_j}{\gamma_j}}{\frac{1}{N}\sum_{j=1}^{N}\frac{1}{\gamma_j}}\right] + \frac{1}{\gamma_i}\left[log\rho_i - \frac{\frac{1}{N}\sum_{j=1}^{N}\frac{log\rho_j}{\gamma_j}}{\frac{1}{N}\sum_{j=1}^{N}\frac{1}{\gamma_j}}\right] t + \frac{1/\gamma_i}{\frac{1}{N}\sum_{j=1}^{N}\frac{1}{\gamma_j}}\overline{c}_{ts} \qquad (6)$$

which shows that the consumption of individual i is a linear function of the average village consumption $\overline{c}_{ts}$ and each parameter has been suitably normalized relative to its village average. Equation (6) shows that individual i's consumption increases linearly 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 $\overline{c}_{ts}$. 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.

2.4 Consumption categories and non-homotheticity

Equation 6 relates to the *total* consumption expenditures c_{its} of household i at time t and state s. As noted by Mace (1991), if consumption preferences are homothetic, the same equation also applies to all k consumption categories since, in this case:

$$c_{itsk} = \alpha_k c_{its} \qquad (7)$$

Hence risk pooling can be tested on any component of consumption via equation (6).

Deviation from optimal consumption allocation may nonetheless arise in practice. In high income countries, for instance, many consumers have insurance policies that cover specific expenditures, such as health care, while other expenditure categories are not insured. In this case, even if total household expenditures c_{its} follow $\overline{c}_{ts}$ one-for-one, we expect health care expenditures to vary less than one-for-one to variation in aggregate expenditures (i.e., to be better insured) than non-health expenditures which, as a result, would have to vary more than one-for-one. Mace (1991) investigated this possibility by testing whether all consumption categories follow equation (6) in US household consumption data. Ignoring heterogeneity in time and risk preferences, she found that non-durable consumption categories follow aggregate consumption $\overline{c}_{ts}$ one-for-one. Under the maintained assumption of homothetic preferences, this indicates that consumption of non-durables is not distorted by good-specific insurance and thus that perfect risk pooling is achieved.

This interpretation may be misleading if consumption preferences are not homothetic, a situation that is particularly likely to arise in our study population where food is a necessity while non-food non-durables have an elasticity above one, i.e., they increase more than proportionally with the expenditure budget of the household. To illustrate how this affects the risk pooling test, imagine a consumer who optimally consumes 70% of her budget on food and the rest on non-food non-durables when her budget is 100, but only 50% on food when her budget is 200. When her total consumption expenditures increase from 100 to 200, her optimal food expenditures go from 70 to 100 while non-food expenditures go from 30 to 100. It follows that a one-for-one doubling of both consumption expenditures would indicate that this consumer is *off* her optimal consumption allocation and thus that her utility is *not* increasing one-for-one with aggregate utility. In other words, it would suggest that this outcome is inefficient and perfect risk pooling is not achieved.

This example demonstrates that if consumption preferences are not homothetic, equation (7) no longer holds. Instead, we have:

$$c_{itsk} = \alpha_k(c_{its}) \tag{8}$$

where $\alpha_k(.)$ denotes an Engel curve. If the shape of this Engel curve is monotonic and can be estimated separately, e.g., from an analysis of the relationship between consumption shares and total consumption expenditures in a cross-section, model (6) can still be fitted to specific consumption categories provided the dependent variable is suitably transformed as:

$$\hat{c}^k_{its} \equiv \hat{\alpha}_k^{-1}(c_{itsk}) \tag{9}$$

where $\hat{\alpha}_k^{-1}(c_{itsk})$ is the inverse Engel curve: it gives the level of total expenditures that is consistent with a given consumption level of good k. It follows that if a household is on its Engel curve, then its total consumption expenditure should be that predicted by equation (9) for each good.[11] Hence if we estimate model (6) separately for different goods, estimated regression coefficients should be identical. If they differ significantly across goods, this implies that households are off their Engel curve. With this approach, we are back in the Mace (1991) world: in an efficient risk pooling economy, applying model (6) separately 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 the consumption preferences of individual households.

It is nonetheless conceivable that risk pooling focuses more on basic necessities such as food, but ignore luxuries. In this case, the consumption share of luxuries should 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. This can be investigated by comparing coefficient estimates of model (6) applied to consumption categories $\hat{c}^k_{its}$ with low and high income elasticities. Examples of policies that provide insurance against fluctuations in food consumptions include school meals for children, the US Food Stamp program (now replaced by SNAP), and the Indian Public Distribution System of shops providing food at subsidized prices for individuals below the poverty line. In a low-income environment such as our study, informal risk sharing may similarly favor stability in food consumption – something we test formally using equation (9).

[11] This approach is reminiscent of the work of Atkin et al. (2023), who propose a methodology to estimate household welfare from consumption of a subset of goods. See also Ligon (2020) who proposes an alternative measure of household welfare using detailed consumption and price data.

2.5 Summary of predictions

We now summarize the main model predictions regarding efficient risk pooling:

1. Individual consumption c_{its} is independent of individual cash-in-hand 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 smooths 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 Engel curve, expenditure shares on specific goods all co-move identically to aggregate village expenditures c_{ts}
7. Liquid assets are used to buffer income shocks at the village level.

The next Section turns these generic predictions into an implementable testing strategy.

3 Testing strategy

Before we present our testing strategy in detail, we must first recognize that, while the model presented in Section 2 applies at the individual level, in our data, as in most, 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 and 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. This implies that, in order to obtain a correct village average $\overline{c}_{ts}$, we must weight each household's per-capita consumption by the number of its members.[12] The same reasoning applies to income and assets, as well as to the risk pooling 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.[13] In practice, we measure the size of each household by its number of adult-equivalents to reflect the fact that consumption needs vary by age and gender. We present in Appendix D a comparison of our results with and without weighting by household size.

3.1 Standard risk pooling test

We test the risk pooling predictions, as listed in section 2.5, under both CARA (in levels) and CRRA (in logs) functional forms. We discuss CARA model test strategy in detail below and derive similar tests under the CRRA model in Appendix C. We start by testing the predictions of the CARA model under the assumption of homogeneous risk and time preferences. Since there is only one realized state of the world per time period, equation (6) simplifies to a perfect risk pooling relationship of the form:

$$c_{it} = \beta_i + \beta_1 t + \beta_2 \overline{c}_t \tag{10}$$

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 Chaudhuri, 1997), we first-difference equation (10) to eliminate the individual specific welfare weight term β_i. We also add two regressors: income y_{it}, which refers to the flow of income collected during period interval t; and cash-in-hand x_{it} which refers to the sum of income and the stock of liquid assets of household i at the beginning of period t (measured in the month of July at the beginning of every panel year or at the end of previous panel year at $t-1$).

The estimated CARA model thus has the form:

$$\Delta c_{it} = \beta_1 + \beta_2 \Delta \overline{c}_t + \beta_3 \Delta y_{it} + \beta_4 \Delta x_{it} + \epsilon_{it} \tag{11}$$

We include cash-in-hand in addition to the income variable that traditionally appears in the efficient risk pooling test so as to allow for the possibility of precautionary saving at the house-

[12]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 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 $66.67 = \frac{1}{3}100 + \frac{2}{3}50$.

[13]To the best of our knowledge, this easy correction is not implemented by Mace (1991); Cochrane (1991) or others. Townsend (1994) mentions the issue in his theoretical model (pp.553-4) but does not weight his regressions by household size.

hold level. Under the null of efficient risk pooling at the village level, the coefficients of both household income and cash-in-hand should be 0 while the coefficient of average village consumption should be 1. In contrast, under a pure household precautionary savings model, household consumption should not vary with village consumption or household income, but it should vary more strongly with household cash-in-hand (Zeldes, 1989a,b) – in which case its coefficient is the so-called marginal propensity to consume. If we find a significant coefficient for income *and* cash-in-hand (in which income is included) – this indicates excess sensitivity to income, e.g., because of asset liquidation costs or mental accounting. Since the household can make up for certain income shocks by engaging in other income earning activities – e.g., redirecting labor supply to the market – the regression model (11) examines the extent to which unsmoothed variation in income is reflected in household consumption. Since cash-in-hand is measured at the beginning of time interval t, it does not include any transfer of funds across households that may have served to smooth period t consumption – which is exactly what we want.

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_2 when the null of perfect risk pooling is false (see Appendix F for illustration).[14] To correct for this bias, we estimate (11) by replacing the village mean $\overline{c}_t$ by the leave-out-mean $\overline{c}_{-i,t} \equiv \frac{1}{N-1}\sum_{j\neq i} c_{jt}$.[15]

Our main null hypothesis is that risk pooling is efficient, which implies that $\beta_2 = 1$ (Prediction 1 in Section 2.5) and $\beta_1 = \beta_3 = \beta_4 = 0$ (Prediction 2 in Section 2.5). Equation (11) 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 $1 > \beta_3 > 0$ and $1 > \beta_4 > 0$

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:

$$\Delta\overline{c}_t = \beta_1 + \beta_3\Delta y_t + \beta_4\Delta x_t + \epsilon_t \tag{12}$$

Efficient precautionary saving requires asset integration, which implies that consumption co-moves with cash-in-hand ($1 > \beta_4 > 0$): when available cash falls, consumption is reduced, but

[14] For instance, when the true $\beta_2 = 0$ in equation (11), the OLS estimate has a bias equal to $1/N$.

[15] It is easy to show that, under the null of perfect risk pooling, estimating (11) with the leave-out-mean still yields the correct estimate of $\hat{\beta}_2 = 1$ but multiplies α_i by $\frac{N}{N-1}$. See Appendix F for details.

less than one-for-one, and consumption should not vary with current income ($\beta_3 = 0$), since it is already included in cash-in-hand. 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 (12) in CRRA log form.

3.2 Consumption categories and Engel curves

Next, we estimate inverse Engel curves (equation 9) 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[16] – 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.[17]

To implement this idea in the simplest way, model (11) 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 (Prediction 6 in Section 2.5). 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.

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

[17] To illustrate with an example, imagine that a household optimally spends Rs.700 on food and Rs.300 on non-food when its total expenditures are Rs.1000, and Rs.800 on food and Rs.400 on non-food when its total expenditures is Rs.1200. Then if this household is unconstrained and we observe it to spend Rs.800 on food, its total expenditures should be Rs.1200. If, at the same time, we observe it consuming Rs.300 on non-food, we would predict that its total expenditures is Rs.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.

3.3 Heterogeneous risk and time preferences

We now introduce heterogeneity in risk and time preferences across households to take into account in our analysis the Predictions 4 and 5 from Section 2.5. It is well known that tests of risk pooling are biased in the presence of heterogeneous risk preferences (e.g., Schulhofer-Wohl, 2011; Mazzocco and Saini, 2012; Chiappori et al., 2014). Ignoring heterogeneous risk preferences can lead to an omitted variable bias in the coefficient on household income y_{it} in equation (11). As a result, the estimated coefficient on income may be too large or biased upwards, potentially leading to spurious rejections of full risk pooling.[18]

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 household 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 a second step, to re-estimate model (11) with household-specific risk and time preferences on annual data.

We follow in the footsteps of Chiappori et al. (2014) and Mazzocco and Saini (2012) and test for heterogeneity in preference parameters under the assumption of perfect risk sharing. The assumption of full risk sharing is important for our estimates. If full risk sharing does not hold, the method may not consistently estimate the true preference parameters. However, Chiappori et al. (2014) show that the bias caused by the violation of this assumption is minimal, and the estimates predominantly identify the true preference parameters.[19] Therefore, while we acknowledge that perfect risk sharing is required for our estimates of the preference parameters to be consistent, the small size of the likely bias means that we can interpret the estimates as largely reflecting the true preference parameters even if perfect risk sharing is rejected.

Formally, in the first step we estimate model (6) 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

[18] These results are worked out in detail in Schulhofer-Wohl (2011) and formally proved in Proposition 1 in Mazzocco and Saini (2012). We discuss the issue of bias in the presence of heterogeneous preferences in more detail in Appendix B.

[19] Chiappori et al. (2014) show that, if full insurance does not hold, then the method rather identifies a weighted linear combination of true risk tolerance and cyclicity of income. The study further performs back-of-the-envelope calculations and shows that, even if full insurance fails, the estimated risk preferences are a combination of at least 91 percent risk tolerance and at most 9 percent 'cyclicity of income'. The latter is defined by the authors (see page 9, equation 10) as the coefficient of a regression in which the log of household income is regressed on a household-specific intercept, a household-specific trend, household-specific monthly dummies, and an average village-specific income shock. Using this approach in our annual data, monthly dummies drop out and we estimate a cyclicity of income coefficient of 0.83. Applying this bias estimate to equation (13) in Chiappori et al. (2014), we see that, if full insurance fails but is 'at least as good as in the United States', our risk tolerance estimates are a mixture of 91 percent true risk tolerance and $0.09 \times 0.83 = 0.075$. For instance, if the true risk tolerance is 2, our estimated risk tolerance would be $0.91 \times 2 + 0.075 = 1.895$. Similar calculations for time preference using equation (11) in Chiappori et al. (2014) and our estimate of the coefficient of the time trend of 0.011 imply that, if full insurance fails, our estimate of $\frac{log \rho_j}{\gamma_j}$ would equal 0.91 times the true value $+0.09 \times 0.011 = 0.01$. For instance, for a household with discount factor 0.95 and risk tolerance 2, the estimated ρ_j would be 0.956.

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 (6) 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:

$$c_{it} = \alpha_i + \theta_i t + \beta_i \overline{c}_{-i,t} + \epsilon_{it} \tag{13}$$

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

$$\beta_i = \frac{1}{\gamma_i} \tag{13A}$$

$$\alpha_i = \beta_i\left[\log\eta_i - \frac{1}{N}\sum_{j=1}^{N}\beta_j\log\eta_j\right] \tag{13B}$$

$$\theta_i = \beta_i\left[\log\rho_i - \frac{1}{N}\sum_{j=1}^{N}\beta_j\log\rho_j\right] \tag{13C}$$

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.[20]

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 risk preferences: only relative preferences can be recovered from (13). 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

[20]The reader may wonder whether the recovered risk coefficient β_i measures the curvature of the instantaneous utility function $U_i(c_{its})$ or that of household i's value function. In the original Townsend (1994) model, there are no assets and, consequently, the two coincide. In a precautionary savings model with liquid assets, the household's intertemporal value function inherits part of the curvature of $U_i(c_{its})$ and the curvature of the value function increases monotonically with the curvature of $U_i(c_{its})$ (e.g., Stokey and Lucas, 1989; Deaton, 1991). Hence the estimated risk tolerance coefficient β_i can be said to be isomorphic to the risk aversion coefficient of $U_i(c_{its})$. This statement nonetheless requires the maintained assumption that all households share the same technology for accumulating liquid wealth. Indeed, it is well known that, the more flexible the liquid wealth technology is, the less curvature the value function has relative to its corresponding instantaneous utility function. For instance, if the liquid wealth technology allows a perfect smoothing of consumption, the value function becomes linear in cash-in-hand, irrespective of how much risk aversion is exhibited by $U_i(c_{its})$.

and relative time preference parameters ρ_i can be recovered using the following formulas:[21]

$$\log \eta_i = \frac{\alpha_i}{\beta_i} + \log \left[\frac{1}{\sum_{j=1}^{N} e^{\frac{\alpha_j}{\beta_j}}} \right] \tag{13D}$$

$$\log \rho_i = \frac{\theta_i}{\beta_i} + \log \left[\frac{1}{\frac{1}{N}\sum_{j=1}^{N} e^{\frac{\theta_j}{\beta_j}}} \right] \tag{13E}$$

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. Therefore, our method of estimating preferences is simple to implement and relies on a simple linear regression, and differs from previous studies in important ways.[22]

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 (11) to allow for household heterogeneity in risk and time preferences as follows:

$$c_{it} = \alpha_i + \theta_i t + \beta_i \overline{c_{-i,t}} + \xi y_{it} + \zeta w_{it} + \epsilon_{it} \tag{14}$$

As before, perfect risk pooling requires that $\xi = 0$ and $\zeta = 0$. To test this prediction, we write (14) 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:[23]

$$\Delta\left(c_{it} - \hat{\beta}_i \overline{c_t}\right) - \hat{\theta}_i = \xi \Delta y_{it} + \zeta \Delta w_{it} + \Delta \epsilon_{it} \tag{15}$$

[21] Let $X = -\frac{1}{N}\sum_{j=1}^{N} \beta_j \log \eta_j$ and $\phi_i = \frac{\alpha_i}{\beta_i}$. From equation (13B) 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 in equation (13D). A similar approach yields the ρ_i formula, except for the division by N which comes from the different normalization rule $\frac{1}{N}\sum_{j=1}^{N} \rho_j = 1$.

[22] Chiappori et al. (2014) and Mazzocco and Saini (2012) consider household-pairs and estimate correlation of consumption between each pair of households in the village. Chiappori et al. (2014) use moment conditions generated by consumption of each pair of distinct households in a village to impute preference parameters of each household by GMM. Mazzocco and Saini (2012) test for efficient risk sharing for household pairs by examining whether a household's consumption is monotonically increasing with the sum of consumption for the pair. Their non-parametric test uses a risk-sharing function that allows for a general class of utility functions with heterogeneous risk preferences. On the other hand, Schulhofer-Wohl (2011) treats risk preferences as nuisance parameters that must be eliminated from the full risk sharing equation. The author uses quasi-fixed effects that controls for household specific trends and household specific effects of aggregate shocks, thereby removing any heterogeneity in preferences.

[23] Although regression model (15) 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 constructed variables only appear in the dependent variable (e.g., Murphy and Topel, 1985).

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 July 2010 to June 2015.[24] 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.[25] Households in each village were randomly selected to represent households in four landholding classes: large, medium, small, and landless.[26] The data collection timeline follows the agricultural cycle in India, beginning from July to June. 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. We also include value of food consumed from home production as part of earned income. 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 cash-in-hand as a measure of household assets. Cash-in-hand is constructed as the sum of liquid wealth, earned income and unearned income. 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 (gold and silver, cars and motorcycles, furniture and farm equipment). Unearned income is the sum of income from government transfers and rental income received from renting out land, household or farm equipment.

[24] 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 a gap between July to October 2009. To maintain consistency, this paper uses data beginning from panel year 2010 until 2014.

[25] The eight states are Andhra Pradesh, Bihar, Gujarat, Jharkhand, Karnataka, Madhya Pradesh, Maharashtra, and Odisha. 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.

[26] We can compare the sample to the 2013 census data for 18 Semi-arid tropic (SAT) villages of the 30 study villages. The sample accounts for 11% of the population in these villages. We find no evidence indicating that wealthier households are underrepresented in the survey.

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 July. 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 capita by dividing them by their adult-equivalent weight.[27]. We winsorize the top and bottom 1% of the data to remove outliers and large measurement errors.

Table 1 presents descriptive statistics for the main variables used in the analysis. Annual consumption expenditures per adult equivalent are on average Rs. 15,937 in 2010 rupees. This is equivalent to 3 US$ per day and per adult-equivalent, based on a purchasing parity rate of 14.6 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 $\hat{\alpha}_k^{-1}(c_{itsk})$. For robustness, we also report results using Engel curves estimated using the other years and using all the years together.

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 (13) for each household, only using monthly data on consumption expenditures. We then use the results to recover estimates of absolute risk aversion

[27] 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

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 of regression (13). We also estimate a CRRA version of these parameters using a model similar to (13), 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 (15). 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.[28] 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 constitutes 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 dispersion 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

[28]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.

consumption levels across households, even if they share the same risk preferences and the same welfare weight.

5.3 Inequality

We also infer welfare weights $\hat{\eta}_i$ from estimated coefficients from regression (13), using the formula presented in Section 3.3. The welfare weights are identified from the household-specific intercept in model (13). 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 $\bar{c}_t$ 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 (13), 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 Main empirical 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 re-estimate 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 3, these estimates are obtained from first-difference regressions in levels, using a pooled panel of all the sample households. In all regressions, 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, we reject full risk pooling since the coefficient of cash-in-hand is statistically significant, and the coefficient on average village expenditure (0.769) is significantly different from 1 at the 1% significance level.[29] This finding is qualitatively similar to earlier studies that used income as the exclusion variable (e.g., Townsend, 1994; Ravallion and Chaudhuri, 1997). The small magnitude of the coefficient on cash-in-hand nonetheless suggests substantial risk pooling: a Rs.100 change in annual cash-in-hand is associated with Rs. 2.1 change in annual consumption, all measured in real 2010 rupees per-adult equivalent. This indicates extensive (even if not perfect) mutual risk pooling within villages.

Based on the model predictions discussed in Section 2.5 and empirical hypotheses in Section 3.1, we reject the pure hand-to-mouth and household precautionary savings models, as the coefficient of earned income is significantly different from 1 and that of average village expenditures is significantly different from zero at the 1% significance level. We can also reject a model in which assets are pooled for risk purposes, but incomes are not, as the coefficient on cash-in-hand is significantly different from zero. We find no evidence of excess sensitivity of consumption to household income. 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.

Next, we examine the implications of risk pooling separately for food and non-food, under homothetic and non-homothetic preferences. Under homothetic preferences, consumption shares are assumed constant with income, which implies that, in the absence of constraints on consump-

[29] Because we do not observe all the households in each village, this coefficient may be affected by measurement error in $\bar{c}_{-i,t}$ due to sampling. In Appendix F 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 $0.769/0.976 = 0.787$. In analysis not shown here to save space, we also estimate the contrast estimator proposed by Suri (2011) and we obtain similar results to those presented here. Our results are thus unlikely to be driven by measurement error in average village expenditures.

tion, each consumption expenditure category should, on average, co-move 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 expenditures would be expected to co-vary less strongly with 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).

In Table 2 column (2), we see that food expenditures vary less with average village expenditures than non-food.[30] We also note that non-food expenditures vary more strongly with variations in household cash-in-hand, suggesting less smoothing of non-food expenditures across individuals. This difference is statistically significant at the 1% level using a test of equality of coefficients of village expenditures between the food and non-food regressions. 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 lead us to expect: not only do food expenditures only fluctuate less with assets shocks, but they also fluctuate less with average village consumption than non-food expenditures. This seems to suggest that food expenditures are better protected from collective shocks than non-food expenditures.

This interpretation, however, can be misleading because it incorrectly assumes that consumption preferences are homothetic. With 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.[31] Using this approach, we can test whether variation in food and non-food expenditures co-moves similarly with 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 coefficient estimates on village average expenditure shown in columns (4) and (5) are less different from each other than the coefficient estimates obtained by assuming homothetic preferences in columns (2) and (3). This confirms that assuming

[30]Furthermore, the small magnitude on the earned income coefficient suggests substantial smoothing of food expenditures against idiosyncratic income: A 100 rupee change in earned income is associated with a change in food consumption by 1.5 rupee. In contrast, using the old VLS data sample, Rosenzweig and Binswanger (1993) find that an equivalent change in income is associated with 7 rupees change in food consumption - about five times more than the new sample. This suggests that consumption smoothing has improved in these villages over time.

[31]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 expenditure 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.

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 continue to observe that non-foods co-moves more to aggregate village shocks than food expenditures, but the difference is no longer statistically significant: a test of equality of the coefficient for village expenditures across the two regressions fails to reject the null of equality with a p-value of 0.367.

Taken together, these findings indicate that food expenditures (which have low income elasticity) are *not* statistically more insulated from village-level shocks than non-food expenditures (which have high income elasticity), contrary to what the homothetic coefficient estimates presented in columns (2) and (3). This implies that the excess smoothing of food expenditures evident in columns (2) and (3) is consistent with an optimal reorganization of consumption towards food when individual total expenditures fall.

Lastly, we examine the implications of risk pooling 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. In footnote 19, we also estimated the possible magnitude of the bias in risk and time preferences that may arise when full risk pooling is rejected, and showed it to be small. The second step relies on equation (15) to estimate the coefficients of cash-in-hand and earned income. These results are presented in Panel B of Table 2. 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 cash-in-hand. But the difference is small in magnitude not statistically significant: a t-test of equality of the cash-in-hand coefficients in Column (1) of Panels A and B yields a p-value of 0.83. 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 the interpretation we derived of the findings under homogeneous time and risk preferences.

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 statistical significance of the reported coefficients nonetheless remains the same.

The results are presented in Table 3. Standard errors are clustered at the village-level in all regressions. A number of observations are lost when taking logs due to zero or negative values in cash-in-hand 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.

The first thing to notice is how similar results in Panel A Table 3 are to those reported in Table 2. The coefficient of average village expenditures in column (1) is significantly below 1– 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 cash-in-hand are associated with a statistically significant variation in consumption expenditures – but the magnitude of this covariation is small: a 100% change in cash-in-hand, for instance, is associated with, on average a 7.6% change in household consumption per capita. Furthermore, as for the CARA model estimates in Table 2, the coefficient on earned income is very small in magnitude and statistically indistinguishable from zero, indicating no excess sensitivity to household income.

Given this, the cash-in-hand coefficient of 0.076 can be interpreted as a reduced-form marginal propensity to consume (MPC) out of *cash-in-hand* and is precisely estimated with a standard error of 0.011.[32] It is slightly higher than the MPC out of *income* estimated to be 0.05 by Blundell et al. (2008) for households in the US using PSID data and is on the lower end of the mean MPC out of income of 0.21 reported in a recent meta-analysis of 246 MPC estimates by Havranek and Sokolova (2020). In addition, our findings corroborate recent papers that find substantial heterogeneity in MPC across households (Lewis et al., 2019; Aguiar et al., 2020; Colarieti et al., 2024). The low MPC found in our study is far lower from the original estimate of 0.5 by Campbell and Mankiw (1989), a value that is interpreted by Crossley et al. (2025) as equivalent to 'hand-to-mouth' consumption. This confirms that most households in VDSA villages engage in consumption smoothing strategies of mutual risk pooling.

Turning to Table 3 columns (2) 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. As for Table 2, we find that the difference between the village expenditure coefficients between columns (4) and (5) is not statistically significant: a t-test of equality of coefficients between the two regressions yields a p-value of 0.162. This indicates that, once we correct for non-homotheticity, food and non-food vary with average village consumption in a way consistent with a within-year utility-maximizing allocation of total expenditures. We also do not find that non-food expenditures co-move more strongly to cash-in-hand than food expenditures: a t-test of equality of the cash-in-hand coefficients between the regressions in columns (4) and (5) yields a p-value of 0.78, further rejecting the idea that there is less risk pooling in non-food consumption than in food consumption.

[32] We interpret the estimated cash-in-hand coefficient in the CRRA model as a 'reduced form' MPC - a measure of the co-movement between consumption and available liquidity rather than a causal effect of income on consumption. This interpretation is analogous to many studies where MPC estimates are obtained from reduced-form regressions that approximate the Euler equation (Havranek and Sokolova, 2020).

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 small reduction in the coefficients of cash-in-hand in Panel B compared to Panel A, but the bias correction is small and statistically insignificant (p-value =0.65). This conclusion extends to food and non-food expenditures, whether we assume homothetic consumption preferences or allow for non-unitary income elasticities of consumption. Overall, we conclude that the results for the CRRA model are similar whether or not we control for heterogeneous preferences. But, as for the CARA model, conclusions on whether food expenditures are smoothed better than non-food depend crucially on whether or not we allow for non-homothetic preferences.

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 cash-in-hand. 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 over time (Prediction 3 in Section 2.5). 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 (12) to estimate the response of total village consumption to village cash-in-hand. Model 12 also includes village income as regressor in order 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. 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 approaches certainty equivalent consumption whereby the response of consumption to an income shock is equal to the discount rate (e.g., Zeldes, 1989a). 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 (12) 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.

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 significant but small variation of village consumption with village cash-in-hand. Based on this coefficient estimate, a Rs.100 fall in village cash-in-hand is associated with only a Rs. 7 fall in total village consumption. This indicates a high level of consumption smoothing through precautionary saving – i.e., close to certainty equivalence. This is further supported by the small and insignificant coefficient on village earned income. Overall, these results suggest that villages optimally draw on liquid assets to smooth aggregate income shocks.

Columns (2) to (5) repeat the same procedure separately for food and non-food expenditures. 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 is not statistically associated with village cash-in-hand while village non-food expenditures co-move positively with it – and that this difference is somewhat stronger when correcting for non-homotheticity in consumption. Neither form of consumption displays excess income sensitivity. Taken together, these findings indicate that village food consumption is smoothed across years while non-food consumption varies with village cash-in-hand.

The CRRA results shown in Panel B are qualitatively similar to those reported in Panel A for a model in first differences in logs. But we find a larger covariation between village consumption and village cash-in-hand. Based on the estimates in Panel B Column (1), a 10% fall in village cash-in-hand is associated with a 2.7% fall in village consumption.[33] Furthermore, if we interact cash-in-hand with year dummies, the estimated coefficient drops to 0.4% for 2012, but it rises to between 3% and 4.5% for the other years. Based on the estimates in Panel B Column (3), a 10% fall in village cash-in-hand is associated on average with a 4.6% fall in village non-food consumption but no significant fall in food consumption. Correcting for non-homotheticity (column 5 of Panel B) reduces the estimated coefficient to 3.1%, consistent with the fact that the consumption elasticity of non-food is higher than that of food.

Taken together, these results are consistent with near-complete intertemporal smoothing of food consumption and with substantial ($\approx$ 69%) smoothing of non-food, relative to variation in cash-in-hand. We however find that, in columns (4) and (5), this difference in cash-in-hand coefficients between food and non-food is not statistically significant, yielding a p-value of 0.206 for the results in Panel B. We also find no evidence of excess sensitivity to earned income, confirming the integration of liquid assets and earned income for the purpose of aggregate consumption smoothing.

To conclude our analysis, we compare the estimated coefficients of cash-in-hand in the individual regressions (Tables 2 and 3) to those found in Table (4). If household liquid assets are *separately* used by individual households to smooth income shocks, as in (Zeldes, 1989a), we should find large coefficients in Tables (2) and (3). We also expect much lower coefficients in Table (4): variation in cash-in-hand across households within a village would explain variation in their household consumption relative to the village average, but since this variation averages out at the village level, the correlation between village consumption and aggregate cash-in-hand should be much smaller. In contrast, if liquid assets held by individual households are de facto

[33]Given that the sample accounts for 11% of the census population of the study villages on average, this coefficient is underestimated due to sampling error because the sample average of cash-in-hand is used in the regression in lieu of the average for the entire village. To investigate the magnitude of the resulting downward bias, we conduct a simulation exercise similar to the spirit of footnote 29 and find that correcting for sampling error would only raise the estimated coefficient to 3.4%, leaving our qualitative conclusions largely unchanged.

pooled, as in equation (4), to face village-level year-to-year income variation, we should find a larger cash-in-hand coefficient in Table (4) than in corresponding Tables (2) and (3). This is indeed what we find: estimated coefficients for cash-in-hand are larger for the village-level regressions. This is most pronounced for total expenditures in the CRRA model, where we obtain a coefficient of 0.27 in the village regression compared to around 0.07 in the individual regressions, a difference that is statistically significant at the 1% level. This difference is also large in magnitude (3.8 times larger), confirming that there is considerable pooling of the liquid assets held by village communities to deal with village shocks.

7 Robustness and mechanisms

We have tested an expanded model of perfect risk sharing that allows for precautionary savings. We find strong evidence that household consumption co-varies with average village consumption but also co-varies positively with household cash-in-hand. These findings remain, albeit in a slightly weaker form, when we also allow for household heterogeneity in risk aversion and discount rate. We also find that, correcting for differences in income elasticity across consumption categories, food and non-food consumption vary equally with average village expenditures, rejecting the hypothesis that food consumption is smoothed preferentially. In this Section, we subject these findings to additional analysis to assess their robustness to possible confounds and to examine possible mechanisms behind our results.

7.1 Robustness checks

We start by comparing in Appendix Tables A3 and A4 the results from Tables 2 and 3 with and without weighting the regressions by household size. This shows that failing to weight for household size yields findings that are qualitatively similar, but different in magnitude: not weighting seems to bias the coefficient of village cash-in-hand upward in all CARA regressions, erroneously suggesting a higher level of risk pooling than what is actually present in the data. This confirms that weighting by household size reduces the risk of incorrectly failing to reject the null of full risk pooling.

To check that our results are not an artifact of the choice of a specific year to estimate inverse Engel curves, we re-estimate columns (4) and (5) of Tables 1 and 2 with Engel curves estimated from each of the data years, as well as using all five years together. The results, presented in Appendix Tables A5 and A6, show small differences in estimated coefficients across the columns of both Tables, but no qualitative change. Importantly, t-tests of equality of coefficients between the food and non-food regressions remain non-significant at 5% significance level in 11 out of 12 cases, the only exception is the CRRA model for the year 2010. This confirms the robustness of our findings to the choice of year with which to estimate Engel curves. Our results are also internally consistent: if, as we find, risk pooling does not distort consumption choices, Engel curves can equivalently be estimated using any or all of the years at our disposal, implying that we should get similar results with any choice of year – which is what we find.

7.2 Sub-caste analysis

So far, we have focused on risk pooling within entire villages. This implicitly assumes that the village is the correct unit of risk pooling. A number of studies have however suggested that, in the context of India, this assumption need not be appropriate: 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 and Saini, 2012; Shrinivas and Fafchamps, 2018).

In this sub-section, we replicate our analysis at the level of sub-caste units within study villages.[34] The testing strategy is identical to that used in Section 3, 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 preference estimates to the level of their sub-caste.

Table 5 contains the results for both the CARA and the CRRA model and 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 cash-in-hand and earned income coefficients which tend to be slightly larger. Taken together, this suggests that risk pooling within sub-castes within each village is less strong than pooling across all households in the village.[35] 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 smaller in Panel B – but not so much smaller as to change the take-away message of the analysis.

Table 6 replicates the precautionary savings regression analysis presented in Table 4, at the level

[34] 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.

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

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 statistically significant but small for a poor population such as the one we study. We find no excess sensitivity to income, as in Table 4.

Taken together, these results suggest that, at the time our data were collected, sub-castes were no longer the main social and economic unit at which risk pooling was taking place.

7.3 Consumption shares and relative prices

We have found that different components of consumption are not smoothed in the same way. Could this be an artifact of changes in relative prices? Our concern is that covariate shocks may induce a correlation between idiosyncratic income and consumption that could lead to false rejection of full risk pooling, as shown by Ligon (2023). For instance, a negative co-variate shock such as a drought may lead to a fall in the price of food relative to non-food while a positive co-variate shock may lead to a rise in relative food price. Such co-variate shocks would result in a reorganization of consumption towards food in bad years and away from food in good years. This in turn would result in a smoother food consumption across time, and a more variable non-food consumption.

To account for changes in relative prices, we would ideally need to control for a village-specific index of the price of food relative to that of non-food. Unfortunately, the VDSA survey data does not include any information about non-food prices. To overcome this difficulty, we construct two separate variables: a village-specific food price index, which we can construct from the data; and a time-varying variable capturing the evolution of the relative price of food versus non-food at the national level. Given that non-food products include manufacturing goods and services that are traded at the national level, the combination of the two variables should proxy for relative price changes, one component of which varies across villages over years, and one component of which is largely determined at the national level.

The village food price index is constructed using the village price data and consumption quantity data, which are reported in the VDSA Monthly price schedule and Transaction schedule, respectively. This produces one food price index per village. For each village, this index is set it to 100 for the year 2011 in which the Engel curves are estimated. We then include in the estimating equations this Village Food CPI as an additional regressor, the coefficient of which could be interpreted in terms of price elasticities. We expect food consumption to increase when this village-specific food price index rises, since food is widely believed to be relatively inelastic. The ratio of national non-food price index to the national aggregate price index is obtained from national CPI data. It measures the price change of non-food relative to food. We expect food consumption to fall with an increase in the relative price of non-food.

Estimation results are reported in Appendix Table A1 and A2 for the CARA and CRRA models respectively. Consistent with our predictions, they show that food consumption decreases with an increase in relative price of non-food. However, the standard errors on these coefficients are large, probably due to the lack of variation in price indices, as reported in the summary statistics Table 1. Nonetheless, these results validate that our main results on full risk pooling from the CARA and CRRA model in Tables 2 and 3 remain unchanged, suggesting that our findings are unlikely to be driven by purely changes in the relative price of food vs non-food.

If changes in relative prices cannot account for differential consumption smoothing for food and non-food, what else could? One possibility is suggested by the observation that public or private social assistance to the poor often seeks to restrict the use that is made of given funds. Examples include food aid and food stamps, food and housing subsidies, shelter for the homeless, and subsidized education and health care. As mentioned in Section 2.4, these forms of aid can constrain individual choices (Cunha, 2014) and they reflect paternalistic preferences on the part of those who give – i.e., voters in the case of government programs, and donors in the case of non-governmental organizations (Currie and Gahvari, 2008). But evidence of such paternalistic altruism had not yet been documented for informal risk sharing.

7.4 Liquid wealth and financial reconciliation

It is of policy interest to ascertain which components of wealth are used to smooth village consumption in response to income variation (Prediction 7 in Section 2.5). To throw some light on the issue, we follow the recent literature on how households adjust different components of liquid assets in response to changes in household income (Krueger et al., 2023; Colarieti et al., 2024).

In an optimal village risk pooling equilibrium, village liquid wealth L_{vt} can compensate a shortfall in village income Y_{vt} by dissaving $D_{vt} \equiv L_{vt} - L_{v,t-1}$. Given the village's accounting identity $C_{vt} = Y_{vt} + L_{v,t-1} - L_{vt}$, estimating a regression of the form:

$$D_{vt} = \alpha_0 + \alpha_1 Y_{vt} + \alpha_2 C_{vt} + u_{vt} \tag{16}$$

should yield an estimated coefficient $\hat{\alpha}_1$ equal to -1. In practice, measurement error biases $\hat{\alpha}_1$ towards 0.

In our study, village wealth has five main components: land a_{vt}; financial assets f_{vt}; jewelry g_{vt}; consumer durables b_{vt}; and livestock l_{vt}. There are hardly any land sales and purchases in our data, which rules out the use of land as buffer stock against income shocks. We therefore define $L_{vt} \equiv f_{vt} + g_{vt} + b_{vt} + l_{vt}$. Of the four components of L_{vt}, financial assets and jewelry are expected to be the most liquid and consumer durables to be the least liquid due to the gap between the buying and resale prices. It is also conceivable that livestock is an asset that farmers are reluctant to part with (e.g., Fafchamps et al., 1998; Kazianga and Udry, 2006).

A detailed discussion of these ideas is presented in Appendix G. From these observations, we expect more adjustment in f_t and g_t and less in b_t and l_t when the village dissaves. This can be tested by estimating regression:

$$D_{vt}^j = \alpha_0^j + \alpha_1^j Y_{vt} + \alpha_2^j C_{vt} + u_{vt}^j \tag{17}$$

for $j \in \{f, g, b, l\}$. If all four assets are equally liquid, we should find $\hat{\alpha}_1^j = s_{vj}\hat{\alpha}_1$ where $s_{vj}$is the share of asset j in L_{vt}. But if, as we expect, the most liquid assets are used more to smooth income shocks, their estimated $\hat{\alpha}_1^j$ should be larger.

We estimate equation (17) for total village liquid wealth and its different components. Results are reported in Appendix Table A7. In the first column we show how total dissaving evolves with income while controlling for consumption. The estimated coefficient of -0.093 is much smaller than what we should mechanically find if identity (16) held exactly. This suggests the presence of considerable measurement error in financial data, a common feature in household surveys (e.g., Lim and Townsend, 1998; Zhang, 2024b). Consequently, we expect the α_1^j estimates to be similarly biased towards zero and less likely to be statistically significant. With this in mind, we find that most of the variation in village assets associated with shortfalls in income is borne by financial assets and by jewelry (i.e., gold and silver). Livestock hardly co-moves and consumer durables have the wrong sign. This suggests that, as predicted, it is the most liquid assets that adjust to village income shocks.

Given these findings, we conduct an alternative reconciliation test using the data on financial *flows* collected in the ICRISAT survey. Households were asked monthly to report inflows and outflows of gifts and remittances, savings and deposits, as well as formal and informal loans.[36] These figures were tallied separately from the questions about end-of-year stocks of financial and non-financial assets that were used in Table A7. The figures from the two sets of questions do not match, another reminder of how challenging the collection of household cash flow data is. We combine these monthly data to construct net annual flows for each asset category, and we regress these annual flows on income and consumption, as in regression (17).

Results are shown in Appendix Table A8. We find that inflows of funds from savings and deposits and from formal lenders have the expected negative sign, and the magnitude of the estimated coefficients is comparable to those reported in Appendix Table A7. In contrast, gifts and remittances and informal loans have a positive coefficient, indicating that they do not serve to fill the village-level gap between consumption and income. These results, as those reported in Appendix Tables (A7), should only be regarded as suggestive, given that the extent of measurement error in the data makes financial reconciliation impossible, in spite of the considerable level of care exercised by the ICRISAT survey team.[37]

[36]Gifts and remittances are those given to and received from friends, relatives and outsiders. Savings and deposits relate to inflows and outflows of cash from fixed deposits, chit funds, life insurances, and pensions. Loans from formal sources tally inflows and outflows of cash to and from commercial banks, co-op financial institutions, and the Grameen bank. Loans from informal sources cover inflows and outflows of cash to and from money lenders, input dealers, friends, and relatives.

[37]Appendix H provides a more detailed discussion of the potential of measurement error in consumption,

8 Conclusion

In this paper we have estimated a test of risk pooling in village economies based on a theoretical model that combines within-year pooling of risk between households with intertemporal precautionary saving at the village level. The merging of these two approaches to consumption smoothing is new to this paper. The methodology we propose is simple to implement and incorporates corrections for heterogeneity in risk aversion and time preferences across households, as well as for non-homotheticity of consumption preferences. While corrections for heterogeneity in risk preferences have been implemented in the past by (Schulhofer-Wohl, 2011), Mazzocco and Saini (2012) and Chiappori et al. (2014), corrections for heterogeneity in time preferences and for non-homotheticity in consumption are new to this paper. The usefulness of the methodology is illustrated in a geographical and cultural context similar to that studied by Townsend (1994) and others, but using a rich dataset that is more contemporary than what has been used in previous papers.

Regarding within-year aggregate risk pooling, we find that: household consumption expenditures co-move strongly with the village average, but statistically less than one for one; and they are mildly but significantly correlated with household cash-in-hand. These findings narrowly reject perfect risk pooling but nonetheless support the view that much risk pooling takes place.[38] Since these findings echo much of the existing literature, they demonstrate that earlier findings were robust not only to heterogeneity in risk aversion, but also to heterogeneity in time preferences, non-homothetic consumption preferences, and the integration of liquid assets into the analysis.

Our approach also produces several novel insights. First, while within-village consumption smoothing is achieved primarily through the pooling of idiosyncratic cash-in-hand variation across households, with little reliance on individual savings, year-to-year consumption smoothing does rely significantly on village liquid wealth. This implies that precautionary saving takes place at the level of the local community. Second, the assets that are used to fill the village gap between consumption and income are financial assets and jewelry (i.e., gold and silver). Livestock hardly varies with income shocks, as in Fafchamps et al. (1998) and Kazianga and Udry (2006), and consumer durables have the wrong sign.

Third, there is proportionally more smoothing of food consumption relative to non-food consumption, a pattern that could be misinterpreted as a deviation from optimal consumption allocation. Once we correct for non-homothetic preferences, however, this difference falls in magnitude and is no longer statistically significant, thereby rejecting the possibility that village risk pooling distorts consumption smoothing towards food, e.g., due to paternalistic considerations. Lastly, household consumption co-moves slightly with cash-in-hand, but we find no excess

income and assets in household survey data

[38] The lack of perfect (i.e., one-for-one) risk pooling may be explained by limited commitment in informal risk sharing arrangements (Coate and Ravallion, 1993). Some partial evidence has been provided for (e.g., Ligon et al., 2002; Fafchamps and Lund, 2003) and against this interpretation (e.g., Weerdt and Fafchamps, 2011), without fully closing the debate.

sensitivity of consumption to earned income. Since the standard test of risk pooling (Townsend, 1994) uses income as exclusion restriction, not cash-in-hand, it could be underpowered. This may potentially affect previous papers that use income as exclusion restriction in a test for full risk pooling.

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. We do, however, find that liquid wealth plays a role in the year-to-year smoothing of village consumption.

References

Aguiar, M. A., M. Bils, and C. Boar (2020, January). Who are the hand-to-mouth? Working Paper 26643, National Bureau of Economic Research.

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

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

Atkin, D., B. Faber, T. Fally, and M. Gonzalez-Navarro (2020, March). Measuring welfare and inequality with incomplete price information. Working Paper 26890, National Bureau of Economic Research.

Atkin, D., B. Faber, T. Fally, and M. Gonzalez-Navarro (2023, 09). Measuring welfare and inequality with incomplete price information*. *The Quarterly Journal of Economics 139*(1), 419–475.

Attanasio, O., A. Kochar, A. Mahajan, and V. Surendra (2023, May). Risk sharing, commitment constraints and self help groups. Working Paper 31245, National Bureau of Economic Research.

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

Buchmann, N., P. Dupas, and R. Ziparo (2025, February). The good wife? reputation dynamics and financial decision-making inside the household. *American Economic Review 115*(2), 525–70.

Campbell, J. Y. and N. G. Mankiw (1989, April). Consumption, income, and interest rates: Reinterpreting the time series evidence. Working Paper 2924, National Bureau of Economic Research.

Carletto, G., M. Tiberti, and A. Zezza (2022). Measure for measure: Comparing survey based estimates of income and consumption for rural households. *World Bank Research Observer 37*(1), 1–38.

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

Coate, S. and M. Ravallion (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.

Colarieti, R., P. Mei, and S. Stantcheva (2024, March). The how and why of household reactions to income shocks. Working Paper 32191, National Bureau of Economic Research.

Crossley, T., P. Fisher, P. Levell, and H. Low (2025). Eliciting the marginal propensity to consume in surveys. Technical report, CEPR Discussion Paper No. 20440.

Cunha, J. M. (2014, April). Testing paternalism: Cash versus in-kind transfers. *American Economic Journal: Applied Economics 6*(2), 195–230.

Currie, J. and F. Gahvari (2008, June). Transfers in cash and in-kind: Theory meets the data. *Journal of Economic Literature 46*(2), 333–83.

De Weerdt, J. and S. Dercon (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, 03). Saving and Income Smoothing in Cote d'Ivoire. *Journal of African Economies 1*(1), 1–24.

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

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

Dubois, P. (2001). Consumption insurance with heterogeneous preferences. Can sharecropping help complete markets ? Economics Working Paper Archive (Toulouse) 25, French Institute for Agronomy Research (INRA), Economics Laboratory in Toulouse (ESR Toulouse).

Dunbar, G. R., A. Lewbel, and K. Pendakur (2013, February). 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., S. Hoderlein, A. Lewbel, O. Linton, and S. Srisuma (2021, October). Nonparametric Euler Equation Identification And Estimation. *Econometric Theory 37*(5), 851–891.

Fafchamps, M. (1992). Solidarity networks in preindustrial societies: Rational peasants with a moral economy. *Economic Development and Cultural Change 41*(1), 147–174.

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

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

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

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

Havranek, T. and A. Sokolova (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. and C. Udry (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.

Kinnan, C. (2022). Distinguishing barriers to insurance in thai villages. *Journal of Human Resources 57*(1), 44–78.

Kinnan, C. and R. Townsend (2012, May). Kinship and financial networks, formal financial access, and risk reduction. *American Economic Review 102*(3), 289–93.

Kochar, A. (1999). Smoothing consumption by smoothing income: Hours-of-work responses to idiosyncratic agricultural shocks in rural India. *Review of Economics and Statistics 81*(1), 50–61.

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

Krueger, D., E. Malkov, and F. Perri (2023, November). How do households respond to income shocks? Working Paper 31894, National Bureau of Economic Research.

Lewis, D. J., D. Melcangi, and L. Pilossoph (2019, November). 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, February). Estimating household welfare from disaggregate expenditures. Working Paper 109, Center for Effective Global Action.

Ligon, E. (2023, July). Risk sharing tests and covariate shocks. Working paper, UC Berkeley: Department of Agricultural and Resource Economics.

Ligon, E., J. P. Thomas, and T. Worrall (2002, 01). Informal Insurance Arrangements with Limited Commitment: Theory and Evidence from Village Economies. *The Review of Economic Studies 69*(1), 209–244.

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

Little, P. D., K. Smith, B. A. Cellarius, D. L. Coppock, and C. B. Barrett (2001). Avoiding disaster: Diversification and risk management among east african herders. *Development and Change 32*, 401–433.

Lybbert, T. J. and J. McPeak (2012). Risk and intertemporal substitution: Livestock portfolios and off-take among kenyan pastoralists. *Journal of Development Economics 97*, 415–426.

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

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

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

Pan, L. (2009). Risk Pooling through Transfers in Rural Ethiopia. *Economic Development and Cultural Change 57*(4), 809–35.

Paxson, C. (1993, February). Consumption and Income Seasonality in Thailand. *Journal of Political Economy 101*(1), 39–72.

Paxson, C. H. (1992). Using weather variability to estimate the response of savings to transitory income in thailand. *The American Economic Review 82*(1), 15–33.

Ravallion, M. and S. Chaudhuri (1997). Risk and insurance in village india: Comment. *Econometrica 65*(1), 171 – 184.

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

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

Rosenzweig, M. R. and K. I. Wolpin (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.

Samphantharak, K. and R. M. Townsend (2018, February). Risk and return in village economies. *American Economic Journal: Microeconomics 10*(1), 1–40.

Sargent, T. J. (1987). *Dynamic Macroeconomic Theory.* Harvard University Press.

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

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

Stokey, N. L. and J. Lucas, Robert E. (1989). *Recursive Methods in Economic Dynamics.* Harvard University Press.

Suri, T. (2011). Estimating the extent of local risk sharing between households. Unpublished.

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

Udry, C. (1990, 09). Credit Markets in Northern Nigeria: Credit as Insurance in a Rural Economy. *The World Bank Economic Review 4*(3), 251–269.

Weerdt, J. D. and M. Fafchamps (2011). Social identity and the formation of health insurance networks. *The Journal of Development Studies 47*(8), 1152–1177.

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

Zeldes, S. P. (1989b, 05). Optimal Consumption with Stochastic Income: Deviations from Certainty Equivalence. *The Quarterly Journal of Economics 104*(2), 275–298.

Zhang, S. (2024a). Asymmetric information, intrahousehold cooperation, and consumption choices: Experimental evidence from kenya. Technical report, Tufts University.

Zhang, S. (2024b). Hidden in plain sight: Asymmetric information and hidden income within the household. Technical report, Tufts University.

Figures and Tables

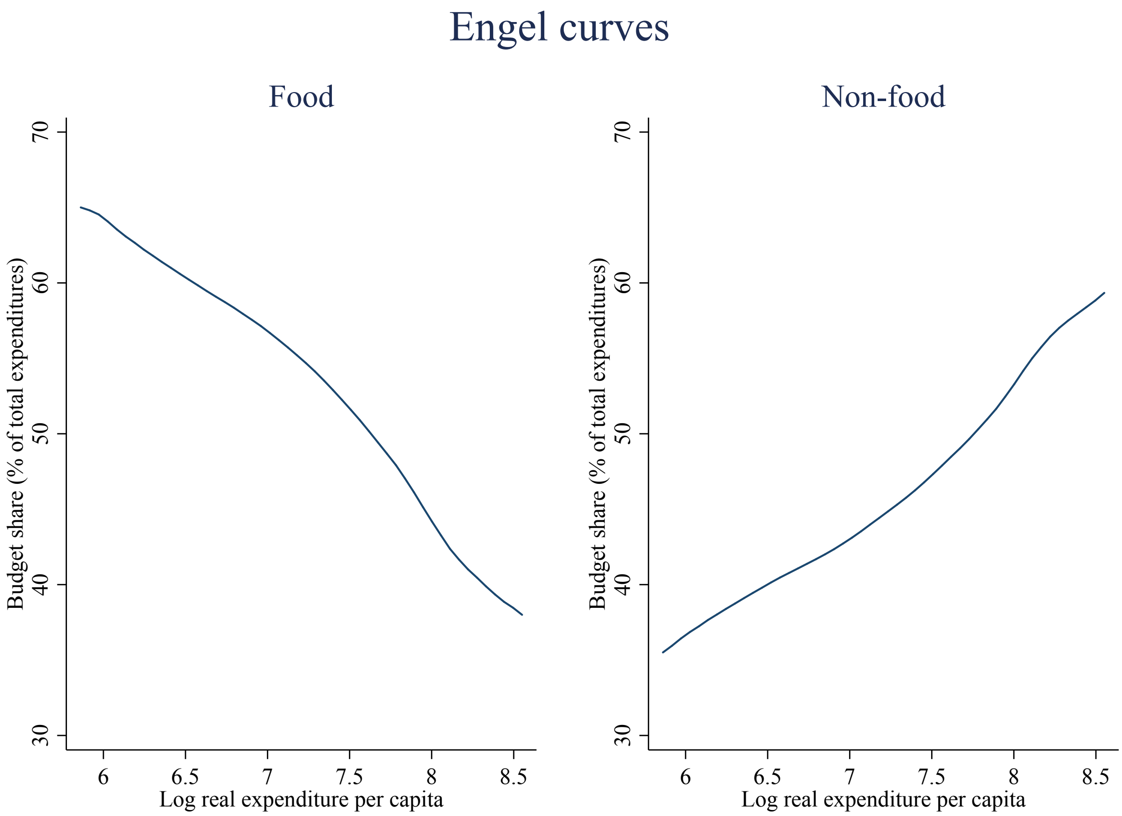

Figure 1: Engel Curves for food and non-food

Engel curves are estimated using cross-section data in 2011, a good rainfall year when village cash-in-hand is the highest in our data. The Engel curves are obtained from a flexible polynomial fit of household budget shares against the log of total household expenditure per capita.

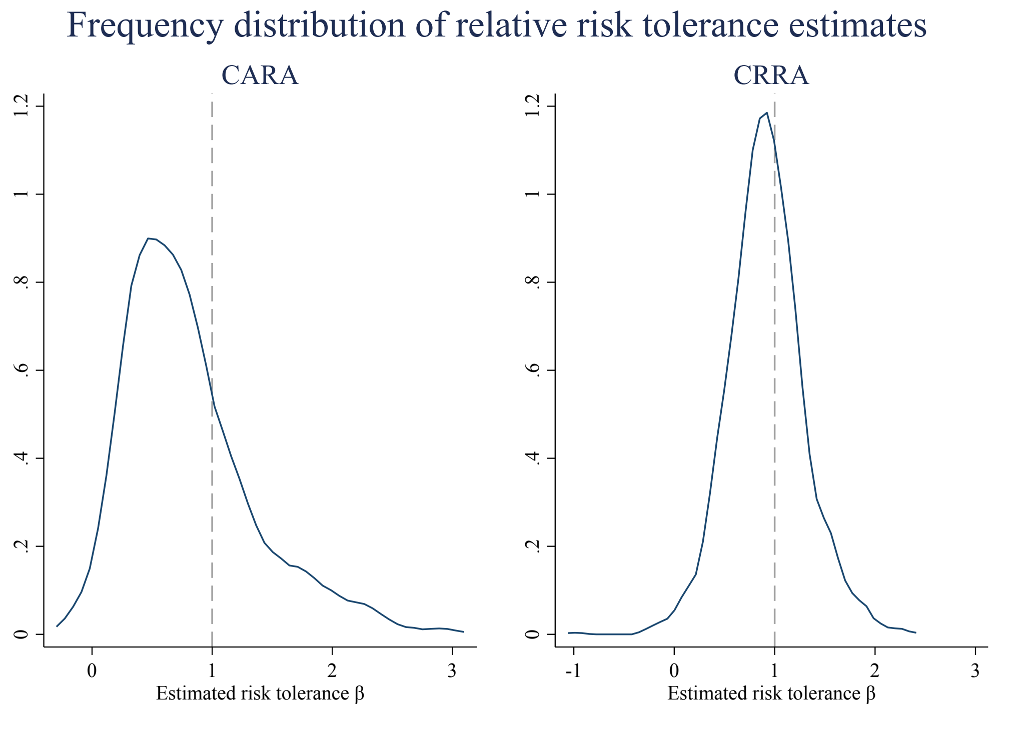

Figure 2: Normalized risk tolerance β

Each panel of this Figure depicts the frequency distribution of $\hat{\beta}_i$, the estimate of household-specific risk tolerance obtained from regression (13). 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.

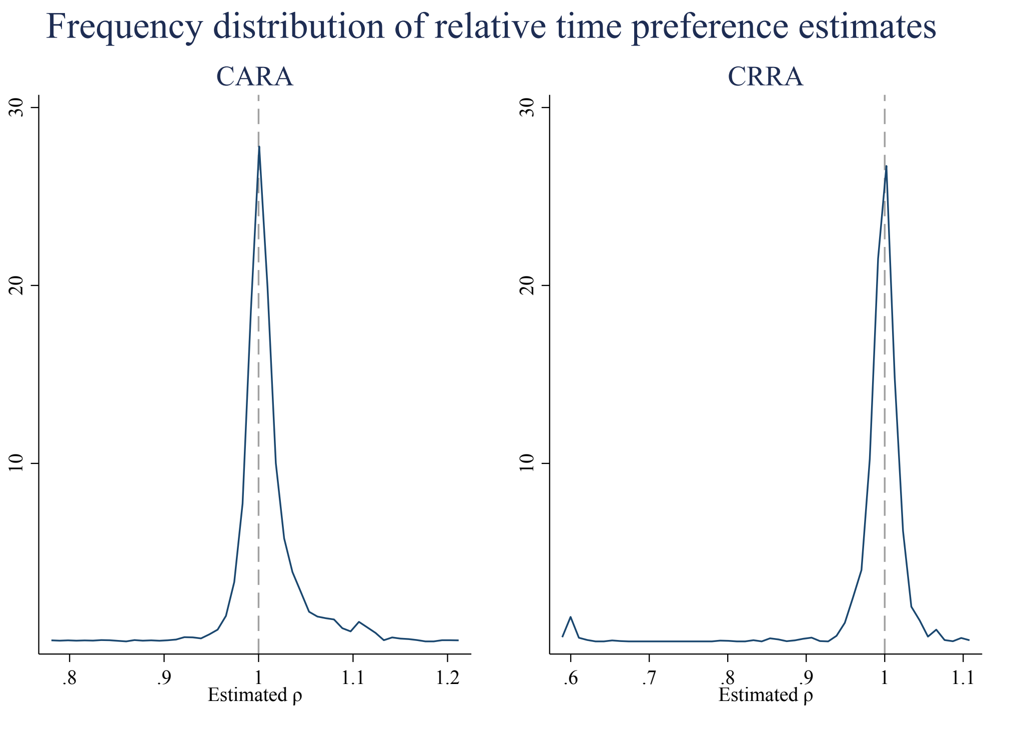

Figure 3: Normalized time preference parameter ρ

Each panel of this Figure depicts the frequency distribution of $\hat{\rho}_i$, the estimate of household-specific discount factor obtained from regression (13). 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.

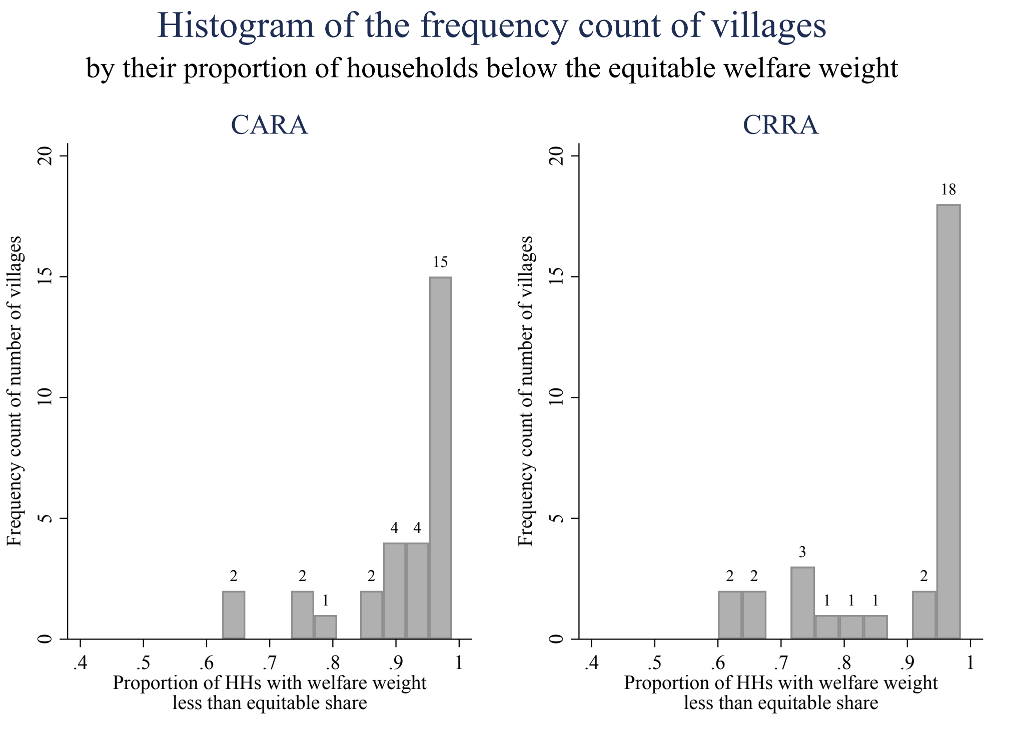

Figure 4: Consumption Inequality

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 (13). 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 households 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	15937	10884
Food expenditure	8321	3776
Non-food expenditure	7615	8072
Income		
Crop and Livestock income	2238	15204
Wages income	15598	20053
Earned income	19774	25100
Unearned Income	2860	7235
Assets		
Wealth	88637	117235
Liquid wealth	25384	44424
Cash-in-Hand	48018	58492
Household characteristics		
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
CPI food village	112	14
Non-food CPI / Agg CPI	0.98	0.02
Households	1296	
Villages	30	
Observations	6589	

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 are 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 eared income. For all values, top and bottom 1% are winsorized to account for measurement error.

Table 2: First differences in levels - CARA model

		Homothetic preferences		Non-homothetic	
	Total expenditure	Food	Non-food	Food	Non-food
	(1)	(2)	(3)	(4)	(5)
Panel A: Homogeneous preferences					
Average village expenditures	0.769***	0.544***	1.014***	0.625***	0.758***
	(0.038)	(0.091)	(0.059)	(0.114)	(0.049)
Cash-in-hand	0.021***	0.012***	0.032**	0.016***	0.020**
	(0.007)	(0.004)	(0.012)	(0.005)	(0.007)
Earned Income	0.013	0.015*	0.010	0.015	0.005
	(0.011)	(0.008)	(0.016)	(0.011)	(0.010)
Observations	5256	5256	5256	5256	5256
Panel B: Heterogeneous preferences					
Cash-in-hand	0.019**	0.007*	0.031**	0.012**	0.017**
	(0.007)	(0.004)	(0.012)	(0.005)	(0.007)
Earned Income	0.010	0.012*	0.007	0.013	0.002
	(0.010)	(0.007)	(0.015)	(0.009)	(0.009)
Observations	5256	5256	5256	5256	5256

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. All regressions are weighted by household-size. 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. Cash-in-hand is the sum of liquid wealth, earned income and unearned income. 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. Top and bottom 1% are winsorized to account for measurement error. 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

	Total expenditure	Homothetic preferences		Non-homothetic	
		Food	Non-food	Food	Non-food
	(1)	(2)	(3)	(4)	(5)
Panel A: Homogeneous preferences					
Average village expenditures	0.895***	0.608***	1.394***	0.653***	0.793***
	(0.020)	(0.064)	(0.134)	(0.067)	(0.074)
Cash in hand	0.076***	0.056***	0.090***	0.062***	0.057***
	(0.011)	(0.012)	(0.023)	(0.013)	(0.013)
Earned Income	-0.001	0.011	-0.018	0.012	-0.009
	(0.007)	(0.009)	(0.012)	(0.009)	(0.008)
Observations	4491	4491	4491	4491	4491
Panel B: Heterogeneous preferences					
Cash in hand	0.069***	0.043***	0.096***	0.050***	0.048***
	(0.011)	(0.012)	(0.023)	(0.012)	(0.014)
Earned Income	0.000	0.012	-0.017	0.013	-0.009
	(0.006)	(0.009)	(0.012)	(0.009)	(0.008)
Observations	4491	4491	4491	4491	4491

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. All regressions are weighted by household size. 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. Cash-in-hand is the sum of liquid wealth, earned income and unearned income. 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. Top and bottom 1% are winsorized to account for measurement error. 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	
	Total expenditure	Food	Non-food	Food	Non-food
	(1)	(2)	(3)	(4)	(5)
Panel A: First differences					
Cash-in-Hand	0.070**	0.008	0.062***	0.015	0.093***
	(0.029)	(0.013)	(0.021)	(0.029)	(0.034)
Earned Income	-0.006	0.018	-0.024	0.043	-0.038
	(0.047)	(0.022)	(0.035)	(0.047)	(0.056)
R-squared	0.084	0.028	0.096	0.028	0.081
Observations	120	120	120	120	120
Panel B: Log differences					
Cash-in-Hand	0.270***	0.117	0.467***	0.122	0.313**
	(0.085)	(0.073)	(0.155)	(0.080)	(0.128)
Earned Income	-0.070	-0.007	-0.162	-0.003	-0.116
	(0.052)	(0.045)	(0.095)	(0.049)	(0.079)
R-squared	0.090	0.035	0.074	0.035	0.049
Observations	120	120	120	120	120

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. Cash-in-hand is the sum of liquid wealth, earned income and unearned income. 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			Homothetic preferences		Non-homothetic	
	Total exp (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										
Average Caste Expenditure	0.545***	0.377***	0.729***	0.450***	0.527***	0.755***	0.498***	1.169***	0.536***	0.676***
	(0.047)	(0.051)	(0.067)	(0.064)	(0.046)	(0.038)	(0.053)	(0.083)	(0.052)	(0.047)
Cash-in-hand	0.022***	0.011***	0.033***	0.016***	0.021***	0.081***	0.057***	0.103***	0.064***	0.064***
	(0.007)	(0.004)	(0.011)	(0.005)	(0.007)	(0.010)	(0.010)	(0.020)	(0.011)	(0.011)
Earned Income	0.012	0.015**	0.009	0.014	0.004	-0.001	0.011	-0.020*	0.013	-0.011
	(0.011)	(0.007)	(0.017)	(0.009)	(0.010)	(0.007)	(0.007)	(0.011)	(0.008)	(0.007)
Observations	4543	4543	4543	4543	4543	3862	3862	3862	3862	3862
Panel B: Heterogeneous preferences										
Cash-in-hand	0.021***	0.010***	0.034***	0.015***	0.020***	0.076***	0.047***	0.108***	0.053***	0.055**
	(0.007)	(0.004)	(0.011)	(0.005)	(0.007)	(0.010)	(0.010)	(0.020)	(0.011)	(0.025)
Earned Income	0.010	0.012*	0.009	0.012	0.002	-0.004	0.008	-0.021*	0.009	-0.009
	(0.010)	(0.006)	(0.017)	(0.009)	(0.010)	(0.006)	(0.007)	(0.014)	(0.008)	(0.009)
Observations	4543	4543	4543	4543	4543	3862	3862	3862	3862	3862

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. Cash-in-hand is the sum of liquid wealth, earned income and unearned income. 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. Top and bottom 1% are winsorized to account for measurement error. 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	
	Total expenditure	Food	Non-food	Food	Non-food
	(1)	(2)	(3)	(4)	(5)
Panel A: First differences					
Cash-in-Hand	0.039***	0.009*	0.031***	0.019*	0.044***
	(0.013)	(0.005)	(0.010)	(0.012)	(0.014)
Earned Income	0.022	0.007	0.015	0.018	0.016
	(0.021)	(0.009)	(0.016)	(0.019)	(0.024)
R-squared	0.065	0.023	0.063	0.025	0.054
Observations	360	360	360	360	360
Panel B: Log differences					
Cash-in-Hand	0.122***	0.099***	0.150**	0.109***	0.094**
	(0.033)	(0.026)	(0.061)	(0.028)	(0.047)
Earned Income	-0.008	-0.017	-0.007	-0.018	-0.005
	(0.019)	(0.015)	(0.035)	(0.016)	(0.027)
R-squared	0.046	0.044	0.023	0.045	0.015
Observations	354	354	354	354	354

*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. Cash-in-hand is the sum of liquid wealth, earned income and unearned income. 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.

Online Appendix

A Additional Figures and Tables

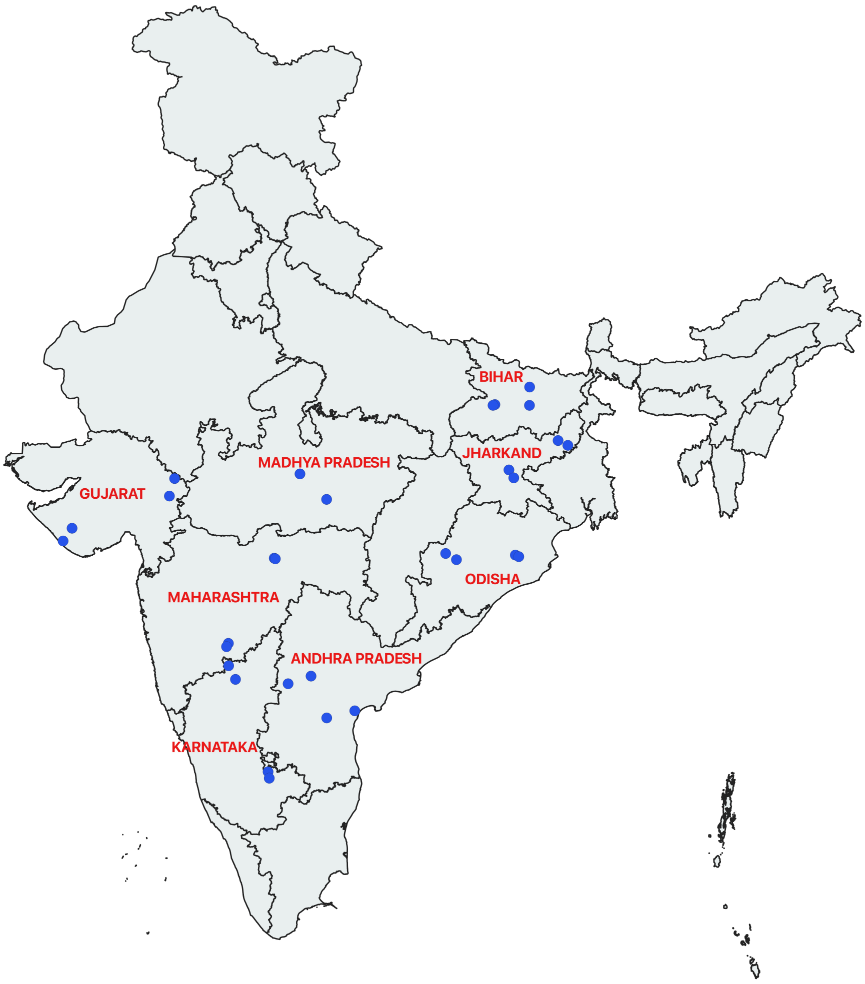

Figure A1: Location of ICRISAT VDSA villages - 30 villages across 8 states in 2010

Table A1: CARA model (difference in levels) controlling for Price index

		Homothetic preferences		Non-homothetic	
	Total expenditure	Food	Non-food	Food	Non-food
	(1)	(2)	(3)	(4)	(5)
Panel A: Homogeneous preferences					
Average village expenditures	0.770***	0.552***	1.007***	0.625***	0.718***
	(0.038)	(0.090)	(0.058)	(0.114)	(0.041)
Cash-in-hand	0.021***	0.012***	0.032**	0.016***	0.019**
	(0.007)	(0.004)	(0.012)	(0.005)	(0.007)
Earned Income	0.013	0.015*	0.010	0.016	0.005
	(0.011)	(0.008)	(0.016)	(0.011)	(0.010)
Village Food CPI	1.665	12.555	-10.234	16.544	-4.502
	(6.814)	(16.868)	(24.754)	(18.910)	(17.381)
Non-food CPI / Aggregate CPI	-9.311	-107.460	97.940	-106.811	68.470
	(32.018)	(100.579)	(81.795)	(119.368)	(56.706)
Observations	5256	5256	5256	5256	5256
Panel B: Heterogeneous preferences					
Cash-in-hand	0.019**	0.008**	0.030**	0.012**	0.016**
	(0.007)	(0.004)	(0.012)	(0.005)	(0.007)
Earned Income	0.010	0.012*	0.007	0.012	0.002
	(0.010)	(0.007)	(0.015)	(0.009)	(0.009)
Village Food CPI	-3.239	11.076	-18.880	13.924	-8.586
	(6.951)	(19.358)	(29.823)	(19.821)	(16.758)
Non-food CPI / Aggregate CPI	-32.651	-155.947*	102.078	-146.922	39.109
	(21.385)	(84.285)	(85.335)	(105.973)	(66.526)
Observations	5256	5256	5256	5256	5256

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. All regressions are weighted by household size. 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. Cash-in-hand is the sum of liquid wealth, earned income and unearned income. Earned income is the sum of income from crop, livestock and wages. Village food price index is constructed using the village price data and consumption quantities, and is set to 100 for the year 2011 in which the Engel curves are estimated. Non-food CPI / Aggregate CPI represents the ratio of national non-food price index to the national aggregate price index from national CPI data. 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. Top and bottom 1% are winsorized to account for measurement error. All standard errors are clustered at the village level. *p<0.10 ** p<0.05 *** p<0.01.

Table A2: CRRA model(difference in logs) controlling for Price index

		Homothetic preferences		Non-homothetic	
	Total expenditure	Food	Non-food	Food	Non-food
	(1)	(2)	(3)	(4)	(5)
Panel A: Homogeneous preferences					
Average village expenditures	0.895***	0.630***	1.350***	0.676***	0.771***
	(0.020)	(0.065)	(0.135)	(0.063)	(0.076)
Cash in hand	0.076***	0.056***	0.091***	0.063***	0.057***
	(0.011)	(0.011)	(0.021)	(0.012)	(0.013)
Earned Income	-0.000	0.012	-0.020	0.013	-0.010
	(0.007)	(0.008)	(0.012)	(0.009)	(0.008)
Village Food CPI	-0.022	-0.063	0.324	-0.050	0.138
	(0.050)	(0.212)	(0.467)	(0.231)	(0.247)
Non-food CPI / Aggregate CPI	-0.011	-0.241*	0.531*	-0.261*	0.257*
	(0.031)	(0.131)	(0.281)	(0.146)	(0.145)
Observations	4491	4491	4491	4491	4491
Panel B: Heterogeneous preferences					
Cash in hand	0.069***	0.043***	0.095***	0.051***	0.048***
	(0.011)	(0.011)	(0.021)	(0.012)	(0.013)
Earned Income	0.001	0.013	-0.019	0.014*	-0.009
	(0.006)	(0.008)	(0.012)	(0.008)	(0.007)
Village Food CPI	-0.052	-0.096	0.300	-0.082	0.107
	(0.044)	(0.244)	(0.529)	(0.253)	(0.229)
Non-food CPI / Aggregate CPI	-0.029	-0.323**	0.622**	-0.332**	0.209
	(0.029)	(0.137)	(0.294)	(0.148)	(0.148)
Observations	4491	4491	4491	4491	4491

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. All regressions are weighted by household size. 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. Cash-in-hand is the sum of liquid wealth, earned income and unearned income. Earned income is the sum of income from crop, livestock and wages. Village food price index is constructed using the village price data and consumption quantities, and is set to 100 for the year 2011 in which the Engel curves are estimated. Non-food CPI / Aggregate CPI represents the ratio of national non-food price index to the national aggregate price index from national CPI data. 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. Top and bottom 1% are winsorized to account for measurement error. All standard errors are clustered at the village level. *p<0.10 ** p<0.05 *** p<0.01.

B Bias caused by heterogeneous risk and time preferences

In this Appendix section, we discuss the possible nature of the bias that heterogeneity of risk and time preferences can induce in the canonical testing regression that does not correct for it.

As discussed in Section 2.3, the theoretical prediction of the full risk pooling model under CARA preferences is that a household's consumption c_{it} varies only with aggregate consumption $\overline{c_t}$, and not with household income y_{it} :

$$c_{it} = \frac{1}{\gamma_i}\left[log\,\eta_i - \frac{\frac{1}{N}\sum_{j=1}^{N}\frac{log\eta_j}{\gamma_j}}{\frac{1}{N}\sum_{j=1}^{N}\frac{1}{\gamma_j}}\right] + \frac{1}{\gamma_i}\left[log\rho_i - \frac{\frac{1}{N}\sum_{j=1}^{N}\frac{log\rho_j}{\gamma_j}}{\frac{1}{N}\sum_{j=1}^{N}\frac{1}{\gamma_j}}\right] t + \frac{1/\gamma_i}{\frac{1}{N}\sum_{j=1}^{N}\frac{1}{\gamma_j}}\overline{c}_t \tag{18}$$

Therefore, the standard reduced-form test of full risk pooling consists of including household income y_{it} as an exclusion restriction parameter:

$$c_{it} = \alpha_i + \theta_i t + \frac{1}{\gamma_i}\overline{c}_t + \xi y_{it} + \epsilon_{it} \tag{19}$$

However, most studies on risk pooling do not estimate (19). Rather, a common co-efficient of risk aversion $\gamma_i = \gamma$ and a common discount factor $\rho_i = \rho$ i.e homogeneous risk and time preferences for all households in the village are assumed and the following is estimated:

$$c_{it} = \alpha_i + \theta t + \frac{1}{\gamma}\overline{c}_t + \xi y_{it} + \epsilon_{it}^{homo} \tag{20}$$

The central point of accounting for heterogeneous preferences is that omitted variable bias makes the estimated coefficient on income ξ in equation (20) too large or positively biased leading to spurious rejections of full risk pooling (Schulhofer-Wohl, 2011; Mazzocco and Saini, 2012)). If the true model is equation (19), but the researcher mistakenly estimates equation (20), then error term in equation (20) is:

$$\epsilon_{it}^{homo} = \left(\frac{1}{\gamma_i} - \frac{1}{\gamma}\right)\overline{c_t} + \epsilon_{it}$$

The coefficient on income ξ in equation (20) is unbiased if $Cov(y_{it}, \epsilon_{it}^{homo}) = 0$, and biased upwards if $Cov(y_{it}, \epsilon_{it}^{homo}) > 0$. Suppose e_i is the elasticity of household i's income to the aggregate shock, then it can be shown that $Cov(y_{it}, \epsilon_{it}^{equal}) > 0$ if $Cov(e_i, \frac{1}{\gamma_i}) > 0$. That is, the income coefficient in (20) is biased upwards if the elasticity of income to aggregate shocks is greater for less risk averse households. Similarly, it can be shown that income coefficient would be biased upwards if income rises more quickly with $\overline{c_t}$ for more patient households. These results are worked out in detail in Schulhofer-Wohl (2011) and formally proved in Proposition 1 in Mazzocco and Saini (2012).

C 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:

$$log\, c_{its} = \frac{log\, \eta_i}{\gamma_i} + \frac{log\, \rho_i}{\gamma_i} t - \frac{1}{\gamma_i} log \lambda_{ts} \tag{21}$$

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 (21). We then obtain:

$$log\, c_{its} = \frac{1}{\gamma_i}\left[log\, \eta_i - \frac{\frac{1}{N}\sum_{j=1}^{N}\frac{log \eta_j}{\gamma_j}}{\frac{1}{N}\sum_{j=1}^{N}\frac{1}{\gamma_j}}\right] + \frac{1}{\gamma_i}\left[log \rho_i - \frac{\frac{1}{N}\sum_{j=1}^{N}\frac{log \rho_j}{\gamma_j}}{\frac{1}{N}\sum_{j=1}^{N}\frac{1}{\gamma_j}}\right] t + \frac{1/\gamma_i}{\frac{1}{N}\sum_{j=1}^{N}\frac{1}{\gamma_j}}\overline{log\, c}_{ts} \tag{22}$$

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

$$\Delta log\, c_{it} = \beta_1 + \beta_2 \Delta \overline{log\, c}_t + \beta_3 \Delta log y_{it} + \beta_4 \Delta log w_{it} + \epsilon_{it} \tag{23}$$

where the two exclusion restrictions in levels present in equation (11) 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:

$$log\, c_{it} = \alpha_i + \theta_i t + \beta_i \overline{log\, c}_t + \epsilon_{it} \tag{24}$$

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

$$\beta_i = \frac{1}{\gamma_i} \tag{24A}$$

$$\alpha_i = \frac{1}{\gamma_i}\left[\log \eta_i - \frac{1}{N}\sum_{j=1}^{N}\frac{\log \eta_j}{\gamma_j}\right] \tag{24B}$$

$$\theta_i = \frac{1}{\gamma_i}\left[\log \rho_i - \frac{1}{N}\sum_{j=1}^{N}\frac{\log \rho_j}{\gamma_j}\right] \tag{24C}$$

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 (13) 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:

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

D Accounting for household-size weights

In this section, we assess the importance of weighting the risk pooling test regressions by household-size. We compare results with and without household weights.

Appendix Table A3 lays side-by-side the results from the CARA model with and without household size weights. Results in Columns (1) to (5) are weighted by household size and replicates Table 2 and results in Columns (6) to (10) are unweighted. Similarly, Appendix Table A4 reports results for the CRRA model with and without household size weights. If we compare the coefficients across weighted and unweighted regression, we can see that in most cases these estimates are slightly higher in magnitude in the unweighted results, with a larger difference for the CARA model than for the CRRA model. In all regressions – except non-food in the homothetic CARA regression – this pushes estimates on village expenditures closer to one, thereby reducing the likelihood of rejecting the null of 1.

For total expenditures in the weighted CARA model in Table A3 Column (1), the joint-test of full risk pooling - that the null of village expenditure coefficient equal to 1 and the coefficient on cash-in-hand equal to zero - is rejected with an F-stat of 18.33 and a p-value of 0.0000, whereas for the unweighted CARA model in Column (6), the test is rejected with an F-stat of 8.62 with a p-value of 0.0012. Although the difference is minimal, these results seem to suggest that we are more likely to reject in the weighted case. We also find similar results under the CRRA model, although the difference is negligible: for the CRRA model in Table A4, the joint test for the unweighted case is rejected with an F-stat of 36.92 with a p-value of 0.0000, compared to a F-stat of 30.17 with a p-value of 0.0000 in the weighted case.

E Robustness to Engel curves estimated using different years

Table A3: CARA Model (first difference in levels) with and without hhsize weights

	WITH Hhsize weights					WITHOUT Hhsize weights				
		Homothetic preferences		Non-homothetic			Homothetic preferences		Non-homothetic	
	Total expenditure (1)	Food (2)	Non-food (3)	Food (4)	Non-food (5)	Total expenditure (6)	Food (7)	Non-food (8)	Food (9)	Non-food (10)
Panel A: Homogeneous preferences										
Average village expenditures	0.769***	0.544***	1.014***	0.625***	0.758***	0.822***	0.585***	1.081***	0.692***	0.787***
	(0.038)	(0.091)	(0.059)	(0.114)	(0.049)	(0.045)	(0.088)	(0.085)	(0.114)	(0.063)
Cash-in-hand	0.021***	0.012***	0.032**	0.016***	0.020**	0.030***	0.014***	0.047**	0.020***	0.028**
	(0.007)	(0.004)	(0.012)	(0.005)	(0.007)	(0.010)	(0.004)	(0.018)	(0.006)	(0.010)
Earned Income	0.013	0.015*	0.010	0.015	0.005	0.009	0.012	0.006	0.013	0.003
	(0.011)	(0.008)	(0.016)	(0.011)	(0.010)	(0.011)	(0.008)	(0.017)	(0.012)	(0.010)
Observations	5256	5256	5256	5256	5256	5256	5256	5256	5256	5256
Panel B: Heterogeneous preferences										
Cash-in-hand	0.019**	0.007*	0.031**	0.012**	0.017**	0.027***	0.010**	0.046**	0.016***	0.025**
	(0.007)	(0.004)	(0.012)	(0.005)	(0.007)	(0.009)	(0.004)	(0.018)	(0.005)	(0.010)
Earned Income	0.010	0.012*	0.007	0.013	0.002	0.006	0.009	0.003	0.010	0.000
	(0.010)	(0.007)	(0.015)	(0.009)	(0.009)	(0.010)	(0.007)	(0.016)	(0.010)	(0.009)
Observations	5256	5256	5256	5256	5256	5256	5256	5256	5256	5256

Notes: This table reports the weighted and unweighted regressions from household panel-pooled estimation of first differences in levels, based on CARA model. Regressions in Columns (1) to (5) are weighted by household-size and Columns (6) to (10) are unweighted. 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. Cash-in-hand is the sum of liquid wealth, earned income and unearned income. 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. Top and bottom 1% are winsorized to account for measurement error. All standard errors are clustered at the village level. *p<0.10 ** p<0.05 *** p<0.01.

Table A4: CRRA Model (first difference in logs) with and without hhsize weights

	Regression with Hhsize weights					Unweighted regressions				
		Homothetic preferences		Non-homothetic			Homothetic preferences		Non-homothetic	
	Total expenditure	Food	Non-food	Food	Non-food	Total expenditure	Food	Non-food	Food	Non-food
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Homogeneous preferences										
Average village expenditures	0.895***	0.608***	1.394***	0.653***	0.793***	0.911***	0.636***	1.379***	0.688***	0.797***
	(0.020)	(0.064)	(0.134)	(0.067)	(0.074)	(0.022)	(0.065)	(0.118)	(0.067)	(0.068)
Cash in hand	0.076***	0.056***	0.090***	0.062***	0.057***	0.076***	0.058***	0.091***	0.065***	0.058***
	(0.011)	(0.012)	(0.023)	(0.013)	(0.013)	(0.011)	(0.013)	(0.018)	(0.014)	(0.011)
Earned Income	-0.001	0.011	-0.018	0.012	-0.009	0.005	0.013	-0.008	0.015	-0.003
	(0.007)	(0.009)	(0.012)	(0.009)	(0.008)	(0.007)	(0.009)	(0.010)	(0.010)	(0.006)
Observations	4491	4491	4491	4491	4491	4491	4491	4491	4491	4491
Panel B: Heterogeneous preferences										
Cash in hand	0.069***	0.043***	0.096***	0.050***	0.048***	0.071***	0.047***	0.097***	0.054***	0.050***
	(0.011)	(0.012)	(0.023)	(0.012)	(0.014)	(0.010)	(0.012)	(0.018)	(0.013)	(0.011)
Earned Income	0.000	0.012	-0.017	0.013	-0.009	0.004	0.013	-0.009	0.014	-0.004
	(0.006)	(0.009)	(0.012)	(0.009)	(0.008)	(0.007)	(0.009)	(0.010)	(0.009)	(0.006)
Observations	4491	4491	4491	4491	4491	4491	4491	4491	4491	4491

Notes: This table reports the weighted and unweighted regressions from household panel-pooled estimation of first differences in logs, based on CRRA model. Regressions in Columns (1) to (5) are weighted by household-size and Columns (6) to (10) are unweighted. 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. Cash-in-hand is the sum of liquid wealth, earned income and unearned income. 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. Top and bottom 1% are winsorized to account for measurement error. All standard errors are clustered at the village level. *p<0.10 ** p<0.05 *** p<0.01.

Table A5: First difference in levels (CARA) accounting for non-homothetic preferences : Robustness to year

	2010		2011†		2012		2013		2014		All Years	
	Food (1)	Non-food (2)	Food (3)	Non-food (4)	Food (5)	Non-food (6)	Food (7)	Non-food (8)	Food (9)	Non-food (10)	Food (11)	Non-food (12)
Panel A: Homogeneous preferences												
Average village expenditures	0.647***	0.737***	0.625***	0.758***	0.626***	0.747***	0.584***	0.719***	0.595***	0.720***	0.616***	0.722***
	(0.136)	(0.044)	(0.114)	(0.049)	(0.113)	(0.043)	(0.115)	(0.040)	(0.106)	(0.041)	(0.116)	(0.041)
Cash in hand	0.017***	0.020**	0.016***	0.020**	0.016***	0.020**	0.014***	0.019**	0.017***	0.019**	0.016***	0.020**
	(0.006)	(0.007)	(0.005)	(0.007)	(0.005)	(0.007)	(0.005)	(0.007)	(0.005)	(0.007)	(0.005)	(0.007)
Earned Income	0.017	0.005	0.015	0.005	0.015	0.005	0.017	0.005	0.012	0.005	0.016	0.005
	(0.012)	(0.010)	(0.011)	(0.010)	(0.011)	(0.010)	(0.010)	(0.010)	(0.011)	(0.010)	(0.011)	(0.010)
Observations	5256	5256	5256	5256	5256	5256	5256	5256	5256	5256	5256	5256
H_0: Village Exp of Food = Non-food (p-value)	0.53		0.28		0.32		0.27		0.27		0.39	
Panel B: Heterogeneous preferences												
Cash in hand	0.014**	0.017**	0.012**	0.017**	0.012**	0.017**	0.010**	0.016**	0.014**	0.016**	0.012**	0.017**
	(0.005)	(0.007)	(0.005)	(0.007)	(0.005)	(0.007)	(0.004)	(0.007)	(0.005)	(0.007)	(0.005)	(0.007)
Earned Income	0.014	0.002	0.013	0.002	0.012	0.002	0.014	0.002	0.009	0.002	0.013	0.002
	(0.010)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.008)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Observations	5256	5256	5256	5256	5256	5256	5256	5256	5256	5256	5256	5256

Notes: † Results presented in the body of the paper. This table reports results on Food and Non-food expenditures under non-homothetic preferences and robustness to the choice of a specific year to estimate Engel curves. Regression are 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. Cash-in-hand is the sum of liquid wealth, earned income and unearned income. 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. Top and bottom 1% are winsorized to account for measurement error. All standard errors are clustered at the village level. *p<0.10 ** p<0.05 *** p<0.01.

Table A6: First difference in log model (CRRA) accounting for non-homothetic preferences : Robustness to year

	2010		2011†		2012		2013		2014		All years	
	Food (1)	Non-food (2)	Food (3)	Non-food (4)	Food (5)	Non-food (6)	Food (7)	Non-food (8)	Food (9)	Non-food (10)	Food (11)	Non-food (12)
Panel A: Homogeneous preferences												
Average village expenditures	0.610***	0.850***	0.653***	0.793***	0.674***	0.869***	0.607***	0.745***	0.692***	0.802***	0.652***	0.793***
	(0.061)	(0.081)	(0.067)	(0.074)	(0.071)	(0.082)	(0.056)	(0.071)	(0.091)	(0.075)	(0.063)	(0.074)
Cash in hand	0.064***	0.056***	0.062***	0.057***	0.063***	0.060***	0.061***	0.056***	0.062***	0.057***	0.063***	0.057***
	(0.012)	(0.014)	(0.013)	(0.013)	(0.013)	(0.015)	(0.012)	(0.013)	(0.013)	(0.014)	(0.013)	(0.013)
Earned Income	0.014	-0.010	0.012	-0.009	0.012	-0.011	0.013	-0.009	0.011	-0.010	0.013	-0.009
	(0.009)	(0.008)	(0.009)	(0.008)	(0.010)	(0.008)	(0.009)	(0.008)	(0.009)	(0.008)	(0.009)	(0.008)
Observations	4491	4491	4491	4491	4491	4491	4491	4491	4491	4491	4491	4491
H_0: Village Exp of Food = Non-food (p-value)	0.018		0.161		0.072		0.127		0.351		0.147	
Panel B: Heterogeneous preferences												
Cash in hand	0.051***	0.049***	0.050***	0.048***	0.051***	0.053***	0.048***	0.046***	0.050***	0.049***	0.050***	0.048***
	(0.013)	(0.014)	(0.012)	(0.014)	(0.013)	(0.015)	(0.012)	(0.013)	(0.013)	(0.014)	(0.013)	(0.014)
Earned Income	0.015	-0.009	0.013	-0.009	0.013	-0.010	0.014	-0.008	0.012	-0.009	0.014	-0.009
	(0.009)	(0.008)	(0.009)	(0.008)	(0.009)	(0.008)	(0.009)	(0.008)	(0.009)	(0.008)	(0.009)	(0.008)
Observations	4491	4491	4491	4491	4491	4491	4491	4491	4491	4491	4491	4491

Notes: † Results presented in the body of the paper. This table reports results on Food and Non-food expenditures under non-homothetic preferences and robustness to the choice of a specific year to estimate Engel curves. Regression are from household panel-pooled estimation of first differences in logs, based on 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. Cash-in-hand is the sum of liquid wealth, earned income and unearned income. 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. Top and bottom 1% are winsorized to account for measurement error. All standard errors are clustered at the village level. *p<0.10 ** p<0.05 *** p<0.01.

F Accounting for Sampling 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:

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

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.

G Village precautionary saving and optimal portfolio

In this Appendix, we examine the optimal durable consumption and portfolio choice that corresponds to the perfect risk pooling hypothesis. To this purpose, we present a stylized model of a hypothetical social planner solving a standard precautionary saving problem. To recall, Y_{vt}represents the aggregate earned income during year t, $C_{vt}$is the aggregate village consumption in that year, $L_{v,t-1}$is the total value of village liquid wealth at the beginning of year t and L_{vt} is the value of wealth at the end of that year. In the absence of measurement error, the following accounting identity must hold:

$$C_{vt} = Y_{vt} + L_{v,t-1} - L_{vt} = Y_{vt} + D_{vt} \tag{25}$$

where dissaving $D_{vt} \equiv L_{v,t-1} - L_{vt}$ and where, for notational simplicity, we have included the return to productive assets into income Y_{vt}. Put in words: consumption must be financed either from income or from dissaving. In the presence of classical measurement error in income, consumption, or wealth, this identity should still be true on average. It follows that, for a given level of consumption C_{vt}, we should observe a negative correlation between income Y_{vt} and dissaving D_{vt}:

$$D_{vt} = \alpha_0 + \alpha_1 Y_{vt} + \alpha_2 C_{vt} + u_t \tag{26}$$

where α_1 should be converging to -1 when measurement error tends to 0. This observation forms the starting point of our testing strategy.

We now turn to the breakdown of wealth across different assets. To derive the efficient risk pooling strategy, we imagine a village social planner who, at time t, optimally allocates village savings L_{vt} across the four categories of assets present in our data: consumer durables b_{vt} (e.g., bicycles, appliances); livestock l_{vt}; gold and silver g_{vt} (typically in the form of jewelry); and the net stock of financial assets f_{vt}(i.e., savings minus outstanding loans due to the rest of the world). By construction, the data always satisfies the following identity:

$$L_{vt} = b_{vt} + l_{vt} + g_{vt} + f_{vt} \tag{27}$$

We are interested in the correlation between Y_{vt} and time changes in each of the components of wealth L_{vt} that can theoretically be expected:

$$D^j_{vt} = \alpha^j_0 + \alpha^j_1 Y_{vt} + \alpha^j_2 C_{vt} + u^j_{vt} \text{ for } j \in \{b, l, g, f\} \tag{28}$$

where D^j_{vt} is the change in asset j between the beginning and the end of the year. This is equivalent to asking whether some components of wealth adjust more to income shortfalls than others. Since all four types of assets are held on average in each village, it cannot be the case that some assets dominate others. Hence, to derive a meaningful demand predictions for all four assets, we make simple but realistic assumptions about what differentiates them.

We start by assuming that all assets can be bought and sold, but that there is a price difference β_j (which we dub 'brokerage fee') between the buying and selling price of each asset type $j \in \{b, l, g, f\}$. For gold and silver and for financial assets, β_j is assumed to be 0 while it is assumed to be moderate for livestock and large for durables. This implies that jewelry and financial assets are the most liquid, while durables are the least liquid. Other things being equal, this predicts that villages should smooth income shocks using jewelry and financial assets first, and durables last (i.e., only when they have exhausted their stock of liquid assets – e.g., Deaton (1991)). Based on these assumptions alone, we expect dissaving in jewelry and financial assets to be the most negatively correlated with income, conditional on a given consumption level C_{vt} – i.e., we expect that α_1^j is more negative for jewelry and financial assets than for livestock, and least negative for durables.

Next, we refine the model to include a flow of consumption utility generated by some assets such as durables and jewelry. For jewelry and the kind of durable goods in our study, the demand for this consumption utility increases more than proportionally with income. It follows that, when income falls, it is optimal for households to shift consumption away from durables and towards necessities such as food products and non-durables. This implies a redistribution of assets away from jewelry and consumption durables. Hence, based on this income effect alone, we would expect that, for a given consumption, income is more negatively correlated with jewelry and durables than with financial assets.This, however ignores brokerage costs which, in the context of our study area, are larger for durables, which lose value over time, than for jewelry, which typically does not – i.e., $\beta_b > \beta_g$. Combining the income and brokerage effects, we therefore expect the negative correlation between income and assets to be strongest for jewelry, weakest for consumer durables, and intermediate for financial assets.

Predictions for livestock are similar to those for financial assets given that, in our data, livestock consumables such as dairy products are imputed a monetary value that is counted as part of income and added to consumption if they are consumed by the household. It is, however, conceivable that livestock production benefits from increasing returns to scale – e.g., due to a minimal herd size to be economically profitable (e.g., Little et al., 2001; Lybbert and McPeak, 2012) – which creates a hurdle in the acquisition of livestock (cow paper showing flypaper effect). There is indeed evidence that livestock-raising households are reluctant to liquidate their livestock in times of duress (e.g., Fafchamps et al., 1998). Based on this, we expect little readjustment of the village asset portfolio away from livestock, and thus a negligible negative correlation between income and livestock. This logic applies, even more strongly, to land – which we have excluded from analysis given the very low frequency of land sales and purchases in our data, and the fact that these transactions take place almost exclusively within the village itself and thus cannot help smooth village consumption.

The predictions for the optimal village risk pooling model can be summarized as follows:

1. In the absence of measurement error, the level of village wealth should be perfectly inversely correlated with income, conditioning on village consumption – i.e., the coefficient α_1 in equation (26) should be -1. Measurement error in income biases the estimated $\hat{\alpha}_1$ coefficient towards 0.

2. In the absence of brokerage costs, increasing returns, and flow of consumption utility from assets, the portfolio composition of village assets should remain constant, conditional on village consumption. Consequently, coefficients α_1^j in equation (28) should be equal to the share s_{vj} of asset j in the village portfolio. Given identity (27), measurement error in income biases $\hat{\alpha}_1^j$ coefficients towards 0 in proportion to their portfolio share, i.e.,

$\hat{\alpha}_1^j = s_j \hat{\alpha}_1$.

3. In the presence of brokerage costs β_j (i.e., a difference between the buying and selling cost of assets), it is optimal for liquid assets (those with lower brokerage cost) to vary more than illiquid assets. Since β_j is lowest for financial assets and jewelry and highest for durables, we expect $\hat{\alpha}_1^f \simeq \hat{\alpha}_1^g < \hat{\alpha}_1^l < \hat{\alpha}_1^b$.

4. In the presence of consumption utility for some assets, we expect portfolio composition to move away from those assets whose consumption is a luxury, such as jewelry and most consumer durables in our data. Combined with brokerage cost effects, this implies $\hat{\alpha}_1^g < \hat{\alpha}_1^f < \hat{\alpha}_1^l$ with the relative position of $\hat{\alpha}_1^b$ depending on the relative strengths of the brokerage cost effect (that makes durables less liquid) and the income elasticity effect (that makes them less consumed when income falls). In practice, we expect the brokerage effect to dominate.

5. While livestock is a relatively liquid asset in the context of our study, the presence of increasing returns in livestock production creates threshold effects that militate towards using livestock as buffer against temporary income shortfalls. In this case, we expect $\hat{\alpha}_1^l \simeq 0$.

The optimal asset portfolio can also vary with asset prices. We do not observe asset price variations in our data, except for jewelry, which we expect to closely follow the international gold price, converted in Indian rupees. [39]Given that gold prices follow more or less a random walk over our period of study, the optimal jewelry share in village wealth should remain constant, which implies selling jewelry when the price of gold goes up.

Table A7: Decomposition of wealth into its components

	Total Dissavings	Components of Dissavings			
		Net Credit [Savings -Dues]	Gold and Silver	Livestock	Other Consumer Durables
	(1)	(2)	(3)	(4)	(5)
Earned income	-0.093	-0.057	-0.049	-0.002	0.022
	(0.092)	(0.055)	(0.053)	(0.019)	(0.040)
Consumption	0.143	0.295***	-0.118	0.010	-0.056
	(0.130)	(0.078)	(0.075)	(0.027)	(0.056)
Observations	120	120	120	120	120
R-squared	0.013	0.117	0.051	0.001	0.009

Notes: This table reports the results from estimating the village accounting identity as described in Section 7.4 at the village-level. Column (1) estimates the response of aggregate dissavings to changes in income, controlling for consumption, and columns (2) to (5) estimate the response of different components of dissavings. 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. Standard errors are reported in parenthesis. *p<0.10 ** p<0.05 *** p<0.01.

[39]While study households also own silver jewelry, gold accounts for most of the jewelry value.

Table A8: Changes in inflows of credit and loans

	Net Inflows			
	Gifts and remittances (1)	Savings and deposits (2)	Loans formal (3)	Loans informal (4)
Earned income	0.016	-0.034**	-0.012	0.004
	(0.011)	(0.017)	(0.014)	(0.019)
Consumption	-0.048***	-0.106***	0.056***	0.067**
	(0.016)	(0.026)	(0.021)	(0.029)
Observations	150	150	150	150
R-squared	0.057	0.205	0.046	0.049

Notes: This table reports the results from regressing annual flows of credit and loans on earned income and consumption as described in Section 7.4 at the village-level. Columns (1) to (4) estimate the response of net inflows of funds from gifts and remittances, savings and deposits, formal loans and informal loans. 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. Standard errors are reported in parenthesis. *p<0.10 ** p<0.05 *** p<0.01.

H Financial reconciliation and measurement error

Lack of complete financial reconciliation in the VDSA data likely reflects measurement error (e.g., Carletto et al., 2022). First, consumption data C_{vt} is not exempt from various sources of measurement error. It is based on answers to retrospective questions which are affected by recall bias. Consumption information is collected through a multiplicity of modules that all have their own recall period, creating the possibility of double counting or missed expenditures. Of particular relevance are family and community events – e.g., weddings, funerals, sport gatherings – that create large expenditures highly concentrated in time. These expenditures are difficult to capture adequately in surveys, but need to be financed somehow. In addition, consumption modules are typically answered by a single person who may not be aware of all the expenditures incurred by other household members – e.g., food away from home, entertainment, transport (e.g., Buchmann et al., 2025). While survey consumption data is an extremely useful measure of material welfare, it remains subject to mismeasurement.

Second, income data Y_{vt}, like consumption, typically comes from retrospective questions and is similarly subject to recall bias. Furthermore, while a large share of the consumption expenditures of poor rural households are often centralized in the hands of one individual, who can answer questions on expenditures, the same does not hold for income. By nature, rural incomes come from many sources – e.g., sales of crops from farm inventories, sales of livestock and dairy products, agricultural day labor, non-farm income. Because these financial inflows accrue to different household members at disparate times, they are difficult to remember and tend to be underestimated (e.g., Zhang, 2024b). Individual incomes may also be misreported as a way of achieving a modicum of financial independence (e.g., Zhang, 2024a). These sources of measurement error at the household level get magnified by sampling error, given that our estimate of village income Y_{vt} is based on a limited sample of survey respondents.[40]

Third, changes in liquid assets $L_{vt} - L_{v,t-1}$ are also subject to measurement error. Financial assets and jewelry are subject to variations in value that do not generate any cash flow, and thus do not contribute to filling the gap between C_{vt} and Y_{vt}. Furthermore, surveyed households may incur new debt without receiving any corresponding monetary inflow.[41] For these reasons, it is customary to find that consumption and income data collected at the rural household level seldom agree with each other, and careful efforts by Lim and Townsend (1998) and Kinnan (2022) to reconstruct asset data and inflows and outflows of cash have proved frustrating.

[40]Sampling bias can reasonably be ruled out, however: using census data from 2011 for 18 of the study villages, we were able to verify that village samples are representative of the village population on all available predictors of income and wealth.

[41]Examples of the first include: debt forgiveness and default penalties, unrealized gains and losses on assets (including interest charges and inflation), and changes in jewelry value due to the prices of gold and silver. Examples of the second include: promises made to the groom's family for the unpaid portion of the dowry, a common practice in India; debts incurred for others or for ceremonial expenditures that are not fully captured by the survey; and debts incurred to cover the purchase of durables or the cost of factors of production.