Institute for Fiscal Studies

Group Size and the Efficiency of Informal Risk Sharing

IFS Working Paper W15/31

Emla Fitzsimons
Bansi Malde
Marcos Vera-Hernandez

Group Size and the Efficiency of Informal Risk Sharing

Emla Fitzsimons, Bansi Malde, Marcos Vera-Hernández*

This Version: October 2015

Abstract

The objective of this paper is to understand and test empirically the relationship between group size and informal risk sharing. Models of informal risk sharing with limited commitment and grim-trigger punishments upon deviation imply that larger groups provide better informal insurance. However, when subgroups of households can credibly deviate, so that sustainable informal arrangements ought to be coalition-proof, the relationship between group size and the amount of insurance is unclear. Building on the framework of Genicot and Ray (2003), we show that this relationship is theoretically ambiguous. We then investigate it empirically using data on the size of the sibships of the household head and spouse in rural Malawi. To identify the relevant potential group within which risk is shared, we exploit a social norm among the main ethnic group in our sample which is such that the brothers of the wife should play a key role in ensuring her household's wellbeing. We find that households in which the wife has many brothers are not well-insured against crop loss events. Importantly, we fail to uncover a similar relationship for the sisters of the wife, ruling out that our findings are driven by wives with many siblings (e.g. brothers) having poorer extended family networks. Calibrating our theoretical framework using values similar to those in our sample produces a relationship between household risk sharing and group size that is similar to that uncovered in the data, indicating that the threat of coalitional deviations can explain our empirical findings.

Key Words: Group size, coalitional deviations, informal risk sharing, extended family networks

JEL Classification: D14, O1, O12

---

[*] Affiliations: Fitzsimons: UCL Institute of Education and Institute for Fiscal Studies (e.fitzsimons@ioe.ac.uk); Malde: Institute for Fiscal Studies and University College London (bansi_m@ifs.org.uk); Vera-Hernández: University College London and Institute for Fiscal Studies (m.vera@ucl.ac.uk). We thank the Mai Mwana team, especially Tambozi Phiri, Andrew Mganga, Nicholas Mbwana, Christopher Kamphinga, Sonia Lewycka, and Mikey Rosato for their advice, useful discussions, and assistance with data collection. We are grateful also to Julia Behrman, Senthuran Bhuvanendra, Lena Lepuschuetz, Carys Roberts and Simon Robertson for excellent research assistance. We thank Orazio Attanasio, Richard Blundell, Antonio Cabrales, Ethan Ligon, Imran Rasul and participants at the IFS work-in-progress seminar, IFS-UCL Phd conference and EDePo Conference for helpful comments and suggestions. We thank Garance Genicot for kindly sharing code for the model of risk sharing with coalition-proof arrangements. Financial support from the ESRC-NCRM Node 'Programme Evaluation for Policy Analysis' Grant ES/I03685X/1 is gratefully acknowledged. Malde also gratefully acknowledges funding from ESRC Future Research Leaders Grant ES/K00123X/1.

1 Introduction

Risk is a salient fact of life in rural areas of developing countries. Moreover, these contexts are characterised by market imperfections such as weak enforcement (also known as limited commitment), costly monitoring, poor infrastructure, and weak government capacity; which lead to missing or incomplete insurance and credit markets, and an absence of government social safety nets.[1] Instead, households rely on a variety of informal mechanisms, such as (informal) transfers and loans from relatives and friends, to deal with the consequences of risk (Besley, 1995). Such mechanisms are usually based on social ties and groups, such as family or friendship, which are typically more effective in overcoming the aforementioned market imperfections (Rosenzweig (1988a), Rosenzweig and Stark (1989), Fafchamps and Lund, 2003, Fafchamps and Gubert, 2007, Angelucci et al., 2015).[2] A sizeable literature finds that these informal mechanisms are remarkably effective in helping households share risk, though they are unable to perfectly protect household wellbeing. Recent, mainly theoretical work, however, suggests that certain features of these groups are likely to influence how effective they are in providing risk sharing (Bloch et al., 2008, Jackson et al., 2012).

This paper aims to study how one important characteristic of informal risk sharing groups – their size (or number of households in the group) – affects the amount of risk sharing they achieve. We establish theoretical predictions and then test these predictions empirically in a setting characterised by almost no formal enforcement mechanisms. Theoretically, in an environment where informal arrangements need to be self-sustaining, two forces are at play in influencing the relationship between group size and risk sharing: on the one hand, when households are sufficiently patient and interactions are repeated, larger groups allow for more diversification of shocks, leading to higher gains from sharing risk. On the other hand, as shown in the seminal paper by Genicot and Ray (2003), when arrangements need to be robust to deviations by sub-groups, larger groups can be destabilised by smaller subgroups that are large enough to provide significant levels of risk sharing, meaning that stable groups that can sustain risk sharing are bounded from the top. This suggests that the relationship between group size and risk sharing is unclear. We extend the set-up of Genicot and Ray (2003) and use simulations to show that the relationship between group size and risk sharing is theoretically ambiguous. Thus, the exact nature of the relationship between group size and risk sharing is an empirical question.

Conceptually, it is important to distinguish between the actual and potential risk sharing group. Empirically, the former poses several challenges: first, it is difficult to measure accurately,[3] and second, it will be endogenous since individuals sort into groups on the basis on unobserved char-

---

[1] A sizeable literature considers the implications of these imperfections on risk sharing: Kocherlakota (1996), Foster and Rosenzweig (2001), Ligon et al. (2003) and Dubois et al. (2008) consider those for the imperfect enforceability of contracts, while Ligon (1998) and Attanasio and Pavoni (2011) study issues related to moral hazard, and Kinnan (2014) highlights the importance of hidden income.

[2] For example, relatives have numerous opportunities to interact with one another, thus reducing the costs of monitoring each others' actions. Moreover, they could use strategies such as shame or even ostracism (both of which are typically not feasible for formal insurance and credit providers to use) to punish renegers in informal arrangements.

[3] For example, self reports are subject to strategic behaviour as shown by Comola and Fafchamps (2015).

acteristics and shocks that are also correlated with risk sharing. To partially overcome this, much prior literature has taken the risk sharing group to be a village (e.g. Townsend, 1994;1995). Though readily observable in a large number of socio-economic datasets, this definition is likely to be too broad, especially since villages can have 500 or more households. We instead focus on the sibship of the household head and spouse, a group that is predetermined.[4] To reflect the fact that not all members of this group will actually share risk amongst each other, in what follows, we refer to it as 'potential group size'.

In the context we study – Mchinji, Malawi – the crucial role of the family for risk sharing has been documented in the anthropology and sociology literatures (Phiri, 1983;Munthali, 2002; Mtika and Doctor, 2002;Peters et al., 2008). This is also reflected in the data we use: 80% of transfers received by a household are from family. Thus, the number of siblings of the head and spouse are a relevant proxy for 'potential group size' in this setting. Moreover, historical well-documented social norms in Mchinji give an important role to the wife's brothers (relative to her sisters) in ensuring her household's wellbeing. Though an individual's sibship size is predetermined, it might still be correlated with unobserved factors that are related with risk sharing. The norms allow us to not only to obtain a more fine grained measure of potential group size, but also provide us with an important dimension of heterogeneity that helps us to allay concerns of such omitted variable bias. In particular, we can build placebo tests using the wife's sisters to ascertain that our findings are not explained by omitted variables associated with larger families.

To investigate the empirical relationship between group size and informal risk sharing, we draw on a rich longitudinal dataset which includes information on household consumption, crop loss incidence (and intensity) and the number of living siblings of the head and spouse (who we refer to interchangeably as husband and wife) to conduct the analysis. We consider how well protected a household's consumption is to idiosyncratic crop losses – an important source of risk in our predominantly agricultural setting – given the size of its extended family. Given the social norms previously mentioned, we define groups separately by relationship to the husband or wife (that is, we consider groups such as brothers of husband, brothers of wife, and so on). The correlation between changes in log household consumption and the incidence (and intensity) of household crop loss provides a measure for risk sharing (see Townsend, 1994; Mace, 1991; and Attanasio and Szekely, 2004, among others).We find that households where the wife has many brothers achieve worse risk sharing in response to crop losses relative to households where the wife has few brothers. A similar, though slightly weaker, pattern is also found for households where the husband has many sisters.

A concern is that these findings could be a result of the fact that households where the wife has many brothers (or husbands have many sisters) are poorer, and therefore more vulnerable to shocks. However, the fact that we fail to find a similar relationship among households where the wife has many sisters, or households where the husband has many brothers alleviates this concern. Of course,

---

[4] A large literature has documented the importance of the extended family for risk sharing in developing countries. See for example, Rosenzweig, 1988a,1988b; Stark and Lucas, 1988; Rosenzweig and Stark, 1989; Foster and Rosenzweig, 2001; Fafchamps and Lund, 2003; Fafchamps and Gubert, 2007; Witoelar, 2013; Angelucci et al., 2015.

such a comparison would form a valid placebo test only if households where the wife (husband) has many sisters (brothers) are similar to those where the wife (husband) has many brothers (sisters). We confirm this is the case, by testing directly for differences in the age, education and ethnicity of the wife (husband) between households where the wife (husband) has many brothers (sisters) and few sisters (brothers). Additional robustness checks indicate that the findings are unlikely to be explained by households with larger numbers of siblings being more vulnerable to crop losses; or by increased competition for production resources (specifically land) among families with many male siblings.

Lastly, we confirm that our empirical findings are compatible with Genicot and Ray (2003). To do so, we calibrate the theoretical model using values (where available) from the data. The calibrated model yields similar patterns between risk sharing and group size as those found in the data, indicating that the threat of coalitional deviations can explain our findings.

The paper contributes to a number of strands of literature: It relates to a small literature investigating the relationship between risk sharing and group size. A number of studies show that the optimal risk sharing groups are likely to be small in the presence of coalitional deviations (Genicot and Ray, 2003, Dubois, 2006 and Chaudhuri et al., 2010) and transaction costs (Murgai et al., 2002). However, when households can choose the risks they face, and have heterogenous risk preferences, larger groups may become stable, as shown theoretically by Wang (2015).

It also relates to the literature investigating risk sharing in the presence of coalitional deviations. Recent contributions have extended theoretically Genicot and Ray (2003) to characterise the optimal risk sharing contract when current transfers can depend on past transfers and shocks (Bold (2009)); and to allow for savings, and the availability of formal and informal risk sharing institutions (Bold and Dercon, 2014). Bold and Dercon (2014) also implement an empirical test of the model using data from funeral insurance groups in Ethiopia. However, they do not consider the relationship between risk sharing and group size.

Finally, the paper contributes to the literature investigating the role of extended families in risk sharing in developing countries. Recent work has documented that market imperfections influence transactions and informal risk sharing arrangements within the family. For example, Foster and Rosenzweig (2001) document that limited commitment, tempered by altruism, is at play in rural India, whileDeWeerdt et al. (2014) show that asymmetry of information among spatially dispersed extended family networks affects interhousehold transfer decisions in rural Tanzania. Baland et al. (2015) document that transfers among siblings in Cameroon follow a system of reciprocal credit, where older siblings support the education of younger siblings, with the expectation that the younger siblings will reciprocate later. [5]. Our analysis complements this literature by considering how the size of extended family networks affects informal risk sharing.

The rest of the paper is structured as follows. Section 2 lays out the conceptual framework, and shows that the relationship between the amount of risk shared and group size is theoretically

---

[5] This literature also finds that social pressure to make transfers among kin leads to less optimal investment decisions, especially for women (Jakiela and Ozier, forthcoming)

ambiguous when coalitions can deviate. Section 3 provides details on the data, and the context, focusing particularly on norms governing extended family relationships in rural Malawi. Section 4 discusses the empirical specification; while Section 5 displays our main results and robustness checks. Section 6 outlines findings of the model calibration. Section 7 concludes.

2 Conceptual Framework

We consider optimal risk sharing in environments subject to imperfect enforceability of contracts. This assumption matches well our empirical setting – rural Malawi – where formal enforcement mechanisms are rarely available. We draw on the set-up in Genicot and Ray (2003), GR hereon, and add to their analysis by considering explicitly (using numerical simulations) the relationship between the extent of risk sharing and group size.

Households are part of a potential risk-sharing group (in our case, the family) of size n. They face a risky endowment, that takes on two values: h or l; $h \geq l$. The probability of drawing an endowment h in any period is π; $0 \leq \pi \leq 1$. Households are ex-ante identical, risk averse and gain utility from consumption. Household utility is increasing, concave and twice-continuously differentiable. There is no storage technology, and neither formal credit nor insurance is available.

To cope with the consequences of risk, households can make and receive transfers following a transfer rule that depends on the number of households in the group that receive the high endowment shock: When a household receives h, and $k-1$ other households also receive h, each household receiving h sends a transfer t_k to a common pool, which is then shared equally among those receiving l. Consumption for households receiving h is thus $h - t_k$, while that for those receiving l is $l + \dfrac{kt_k}{n-k}$.[6] Households observe the endowments, consumption and transfers made and received by all other households in the group. However, this setting is subject to the imperfect enforceability of contracts. Thus, the transfer arrangement needs to be self-sustaining. In particular, it needs to be such that no individual or sub-group wants to deviate from the arrangement, i.e. it should be coalition-proof. The specific definition of coalition-proofness is as in Bernheim et al. (1987), which places a further restriction that sub-groups that deviate should themselves be robust to further deviations. Thus, arrangements need to be self-sustaining to deviations that are themselves credible.

Given the transfer rule, and the coalition-proofness condition, and focusing on stationary arrangements, the optimal risk sharing arrangement (i.e. transfer in each state) can be recovered from the solution to the following optimisation problem (expressed in per-period terms):

$$\underset{t_k}{max}\ v(\boldsymbol{t}, n) = p^n u(h) + (1-p)^n u(l) + \sum_{k=1}^{n-1} p(k,n) \left[\frac{k}{n} u(h - t_k) + \frac{n-k}{n} u(l + \frac{kt_k}{n-k}) \right] \tag{1}$$

subject to

---

[6] Note that the transfer rule makes use of the fact that the group-level aggregate budget constraint for each period must be satisfied.

$$(1-\delta)u(h-t_k)+\delta v(\boldsymbol{t},n)\geq(1-\delta)u(h)+\delta v^*(s)\quad\forall\, s\leq k \tag{2}$$

where δ is the discount factor, and $v^*(s)$ is the per-period expected utility a household could get by deviating to a stable sub-group of size s, and sharing risk in this sub-group in all subsequent periods. The incentive compatibility constraints in Equation (2) imply that the transfer arrangement should be such that the per-period discounted utility for households that achieve a good shock in the current period and make a transfer t_k to the common pool, and expect to achieve future expected utility of $v(\boldsymbol{t},n)$ is greater than the utility it can achieve from deviating in a sub-group s where it consumes its endowment h this period and shares risk with the sub-group s in the future thus attaining an expected future utility of $v^*(s)$.[7]

When no incentive compatibility constraint binds, the first-best allocation, which equalises consumption for all households within the group for each state of the world, is achieved. By contrast, in autarky, when no risk sharing occurs, households consume their own endowment in each period, achieving a per-period expected utility of $pu(h)+(1-p)u(l)$.

Based on this set-up, GR show that a stable risk sharing arrangement may fail to exist for many group sizes, even for high values of the discount factor.[8] Moreover, they show that the size of stable risk sharing groups is bounded from above: essentially, large groups are not stable in the presence of coalitional deviations, since households receiving a good shock can deviate to form sub-groups within which they can still benefit from group-based insurance in the future. Thus, in larger groups, the outside option may potentially be better than in smaller groups (depending on the sizes of possible stable sub-coalitions). Thus, the transfer made by those receiving h will be lower than in arrangements sustained by ostracising a deviator to autarky in the future. This is because those receiving h need to be induced to remain in the group rather than deviate to a sub-group, which could provide higher utility than autarky. In some cases, no postive transfer may exist, leading to the non-existence of a stable risk sharing arrangement.[9]

Our contribution, relative to GR, is to show within the same set-up that the relationship between the amount of risk sharing and group size is ambiguous. The fact that a stable arrangement may not exist for many group sizes, complicates this exercise.[10] In particular, it is not possible to study this analytically. We instead use numerical simulations to shed light on the relationship.

We need to take a stand on how risk is shared in groups of size n where no stable risk sharing

---

[7] Note that this formulation assumes that in the period that an individual deviates, he consumes his endowment, regardless of the sub-group he deviates with; and shares risk with members of the subgroup in subsequent periods.

[8] In models where the risk sharing arrangement is sustained by ostracising individuals who deviate (i.e. deviating individuals revert to autarky in future periods), a stable arrangement may fail to exist when the discount factor is low. When arrangements need to be coalition-proof, however, a stable arrangement may fail to exist even if the discount factor is sufficiently high.

[9] When arrangements can be non-stationary, a larger group could be stable. This is because only a sub-set, rather than all, of potential deviators need to be compensated to remain in the risk sharing arrangement. Nonetheless, GR show that the size of the largest stable group will still be bounded from the top (though it could be larger than the largest stable group under stationary arrangements).

[10] Moreover, as indicated by GR, the existence or not of a stable arrangement for groups of size greater than 2 is sensitive to parameter values.

arrangement exists. One possibility is that households remain in autarky. However, this is not very satisfactory, especially since within this set-up, households can deviate from an autarky punishment by cooperating with subgroups of households. Thus, given that households are ex-ante identical in this setting, a natural assumption is that in cases where no stable arrangement exists for a group of size n, the group randomly partitions into stable subgroups in a manner so as to maximise the sum of expected utility,

$$\sum_{i=1}^{n}\sum_{s\in S} n_s * s * v_i(\boldsymbol{t}, s) \tag{3}$$

where S is the set of stable coalitions (or groups), and i indexes households in the group.[11] In other words, we assume that there exists a social planner who chooses a combination of stable sub-groups such that the sum of expected utility (as in Equation (3)) is maximised, and then randomly sorts households into these sub-groups.[12,13] We can then calculate the expected utility of a household in the unstable potential group of size n as the weighted average of the expected utilities associated with the combination of stable subgroups (the actual risk sharing group) that maximises the potential group's expected utility, with weights calculated as the probability of being randomly assigned to a particular sub-group.

We evaluate the extent of risk sharing using two measures:[14]

- The household's weighted average expected utility,

  $$\sum_{s\in S: s \text{ stable}} \pi_s v_i(\boldsymbol{t}, s) \tag{4}$$

  where π_s is the probability of being in the stable group s. This is the social planner's objective function. The value of this function increases as fluctuations in a household's consumption fall: a larger stable group will have a higher value of $v_i(\boldsymbol{t}, s)$ since (i) the probability of states where all households receive the same shock falls with group size, and so there is more scope for risk sharing; and (ii) households have concave utility.

- The weighted average expected difference in marginal utility between the two endowment realisations,

  $$\sum_{s\in S: s \text{ stable}} \pi_s \sum_{k=1}^{s-1} p(k,s)\frac{k}{s}[u'(c^l_{k,s}) - u'(c^h_{k,s})] \tag{5}$$

---

[11] This need not be the only way by which the group partitions, particularly when households are allowed to be heterogenous. For example, partitions could emerge endogenously as in Ambrus et al. (2014), who allow for different transfers to be made between pairs of households embedded in a network.

[12] In doing so, we assume that unstable groups are arranging themselves in a manner so as to generate the highest possible insurance for their members.

[13] Since households are ex-ante identical, we assume that the social planner places equal weight on each household when deciding how to allocate households in unstable groups to stable subgroups. However, this assumption can be relaxed easily to allow for arbitrary planner weights. However, note that the transfer rule, and thus expected utility, $v_i(\boldsymbol{t}, s)$, will be the same for all households.

[14] The measure used in the empirical analysis is slightly different and is based on ratios of the marginal utility of consumption.

> This measure captures the difference in marginal utility that a household expects between states where it receives h and those where it receives l. In a state where k households receive h, higher values of t_k (upto the value equating $c^l_{k,s}$ and $c^h_{k,s}$) will reduce the gap between $c^l_{k,s}$ and $c^h_{k,s}$, and so reduce the difference $u'(c^l_{k,s}) - u'(c^h_{k,s})$. If transfers are large enough such that $c^l_{k,s} = c^h_{k,s}$ for all states, perfect risk sharing is achieved and this measure will be 0. However, deviations from perfect risk sharing in any state of the world, in any of the stable sub-groups that the group can partition into, would lead to this measure being positive. Moreover, the greater the deviation from perfect risk sharing (i.e. the higher the gap between $c^l_{k,s}$ and $c^h_{k,s}$), the higher the value of this measure. Thus, lower values of this measure indicate better risk sharing.

We next use this set-up to assess the relationship between the extent of risk sharing, as measured by the expressions (4) and (5), and potential group size.

Simulations To simulate the model, we make some assumptions on the functional form of the utility function, and on parameter values. In the examples we show here, we use the same parameter values as in GR (Example 2).[15] Utility is assumed to be of the constant relative risk aversion form, i.e.

$$u(c) \quad = \quad \frac{c^{(1-\rho)} - 1}{(1-\rho)}$$

where ρ is the coefficient of relative risk aversion. n is assumed to be 8, which matches the largest group size in our data (see Section 3 below). ρ is assumed to be 1.6, $\delta = 0.83$, $h = 3$ and $l = 2$ as in GR. Finally, the probability of receiving the high endowment, $p = 0.4$. With this set of parameter values, only sub-coalitions of size 1, 2 and 3 are stable, as reported in GR and documented in the Table 1.

Given this set of stable sub-groups, we compute the two measures outlined in Equations (4) and (5) to evaluate the extent of risk sharing for each of the different potential group sizes. These are plotted in the left and right panels of Figure 1.[16] Weighted expected utility increases with group size for potential groups up to size 3 (which is expected as 3 is the largest stable group), before fluctuating in a zig-zag pattern. The fall with group size is a result of a breakdown in informal risk sharing: in a potential group of size 4, one household would be in autarky, while the other three households could cooperate together and benefit from risk sharing opportunities. The subsequent zig-zag style pattern arises from the combination of stable group sizes that is viable in larger unstable potential groups. A similar picture emerges for the second measure – the weighted

---

[15] We use these parameter values so as to illustrate what happens to the extent of risk sharing in a documented case where only a small number of potential group sizes is stable. In Section 6, we illustrate the patterns of risk sharing and potential group size that emerge when we set the parameter values to match our data.

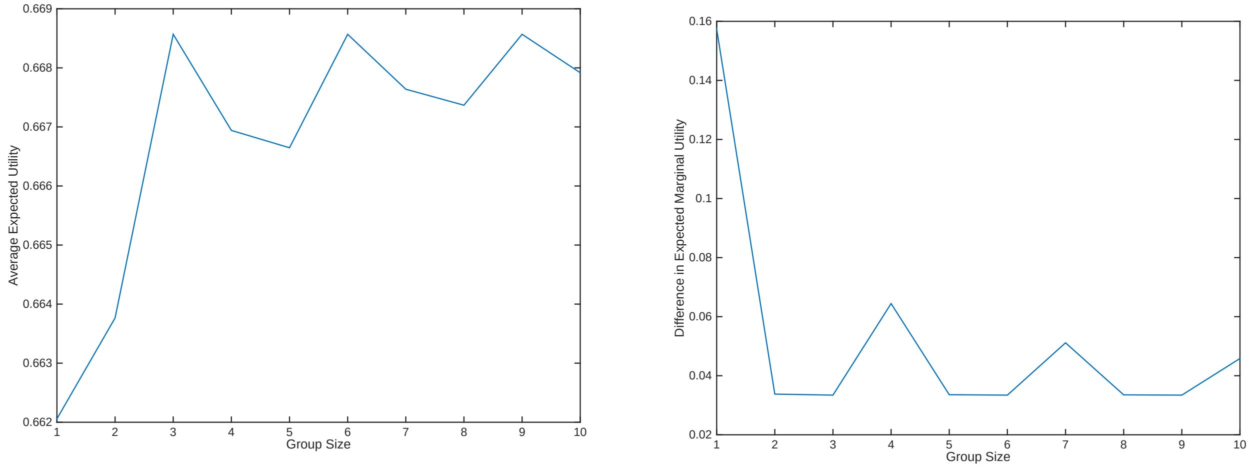
[16] A detailed overview of the calculations that yield the Figure is in Appendix A.

Table 1: Stable groups

Group Size	Parameter Set A
1	✓
2	✓
3	✓
4	×
5	×
6	×
7	×
8	×
9	×
10	×

average expected difference in marginal utility – (right panel, Figure 1), though the pattern is inverted since improvements in risk sharing are associated with decreases in this measure.

Figure 1: Risk Sharing and Group Size - example from Genicot and Ray (2003)



Notes: The Figure on the left panel shows the relationship between weighted average expected utility and group size, while that on the right panel shows the relationship between the weighted average expected difference in marginal utility and group size

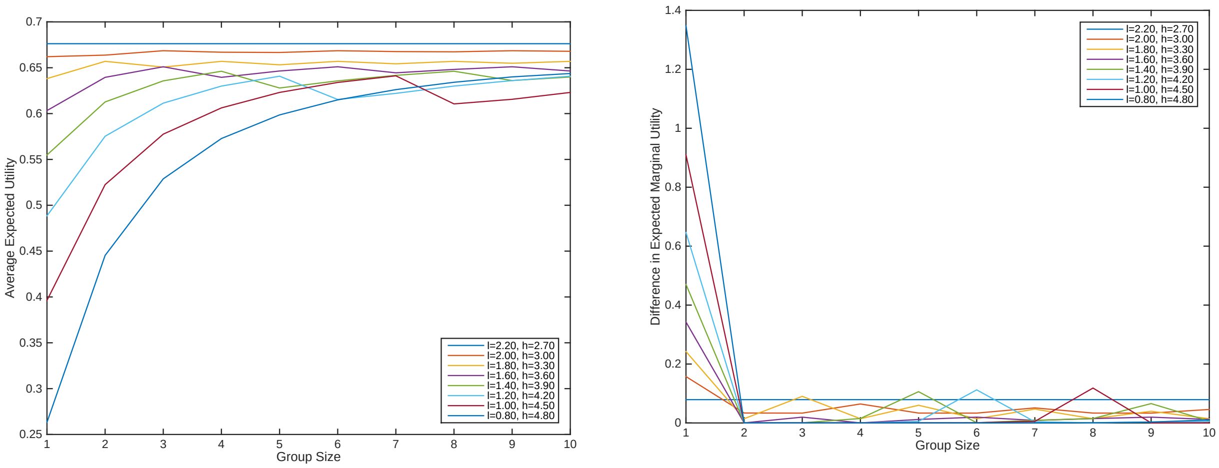
To be noted, though, and documented by GR, is that the stability of groups is sensitive to parameter values. In particular, changing the parameters ρ, p, h and l a little can change which group sizes are stable.[17] This is displayed in the Figure 2, which plots the two measures of the degree of risk sharing for different levels of h and l. The values of these variables have been selected so as to have the same average endowment, but different variances. A higher variance implies a greater need for insurance. The Figure indicates that as the need for insurance increases, larger potential groups become stable, and these groups achieve better risk sharing than smaller potential groups. This is best displayed by the line corresponding with the highest need for insurance ($l = 0.8$; $h = 4.8$),

---

[17] From the repeated games literature, it is well known that groups of size 2 can be unstable for low levels of the discount factor, δ. The instability noted here for larger groups arises even when δ is high.

and is the lowest line in the left panel of Figure 2. This line is increasing monotonically, indicating that all group sizes are stable. By contrast, when the need for insurance is low ($l = 2.2$; $h = 2.7$), a case depicted by the top-most line in the left panel of Figure 2, no potential group of size > 1 is stable.[18]

Figure 2: Risk Sharing and Group Size - Example 2



Notes: The Figure on the left panel shows the relationship between weighted average expected utility and group size, while that on the right panel shows the relationship between the weighted average expected difference in marginal utility and group size

Thus, the simulations indicate even with a small set of parameter values, that there is a theoretically ambiguous relationship between group size and the extent of risk sharing in this model.[19] The nature of this relationship is thus an empirical question, which we now turn to.

3 Context and Data

Our empirical setting is Malawi, one of the poorest countries in Sub-Saharan Africa, with around three quarters of its population living on less than \$1.25 a day. Over 80% of its population lives in rural areas, with subsistence agriculture providing the main source of income for a substantial proportion. Infrastructure in rural areas is very weak, with just one in sixteen households having access to electricity, and one in five households having access to piped water.[20] The main crops grown are maize, tobacco and ground nuts. Agriculture is mainly rain-fed, and agricultural production and income are thus highly dependent on unpredictable weather. Access to formal insurance

---

[18] Average expected utility is nonetheless higher in this case (even in autarky) since the variance of the endowment is much lower in this case.

[19] We note that other models might also imply that the size of the optimal risk sharing group is smaller than the whole potential group. The presence of coordination costs that are increasing in group size could also yield a similar pattern, as shown by Murgai et al. (2002). However, to our knowledge, no work has characterised the relationship between the extent of risk sharing and group size.

[20] Source: Malawi Population and Housing Census (2008).

and financial products and services is low, with only 3% of adults holding an insurance product and less than 20% a formal bank account.[21] Instead social connections, particularly family, are important for providing risk sharing, as we show below.

3.1 Data Description and Sample Selection

We use data from the Mai Mwana - IFS Economic Survey, a longitudinal survey collected in collaboration with the authors in Mchinji District to evaluate two randomised health interventions – a volunteer infant feeding counselling intervention and a women's group intervention.[22] The survey interviewed approximately 3000 women aged 17-43 and their households living in approximately 600 villages across the district. It collected detailed information on household consumption, adverse events, individual labour supply, health indicators, assets and demographics, and importantly for us, information on extended family networks within and outside the village. Two waves of data were collected, in 2008-09 and 2009-10. The panel dimension allows us to better control for household-level unobserved variables that are correlated with our measure of potential group size, crop losses, and risk sharing.

We restrict the analysis to the following sample: (i) Households living in control areas. (ii) Households where the main respondent was resident in the same village over both surveys. (iii) Households where the main respondent in our survey was either the head or the spouse. (iv) Villages with more than 1 household surveyed. Restriction (i) is imposed since the interventions could have altered risk sharing arrangements within the village, by for instance, altering social interactions or improving community cooperation (particularly in the case of the women's groups).[23] Restriction (ii) is imposed to allow us to correctly account for village-level aggregate shocks.[24] We impose restriction (iii) to ensure that we are studying the networks of individuals with relatively similar intrahousehold bargaining power in the sample. Finally, (iv) is imposed because we control for village fixed effects.

Table 2 displays some descriptive statistics of our analysis sample. It contains approximately 524 households living in 102 villages. Note that throughout what follows, we recode the male member of a couple (where available) to be the head, while the female member is designated to be the spouse. A note on terminology is in order: throughout the paper, we will use head and spouse interchangeably with husband and wife. Both the head (husband) and spouse (wife) have low levels of education on average, with approximately 16% (7.4%) of husbands (wives) having some secondary schooling. Further, husbands are older than their wives by on average around 5 years. Households have on average just over 5 members, and most own their own dwelling and land. Despite this, households

---

[21] Source: Finscope Malawi (2009).

[22] See Lewycka et al. (2013) and Fitzsimons et al. (2014) for findings of the impact evaluation. The data is publicly available at http://discover.ukdataservice.ac.uk/catalogue?sn=6996

[23] Fitzsimons et al. (2013) find suggestive evidence of this.

[24] Around 18% of the survey main respondents in the data migrated to another village between 2008-09 and 2009-10. The primary reason for migration was marriage. In additional analysis, we checked whether migration was systematically related with the crop loss, and found no evidence of this.

Table 2: Sample Descriptives

Variable	**N**	**Mean**	**Std. Dev.**
Husband has no education (yes=1)	477	0.140	0.348
Husband has some primary (yes=1)	477	0.222	0.416
Husband has completed primary (yes=1)	477	0.478	0.500
Husband has at least some secondary (yes=1)	477	0.159	0.366
Husband's years of education	477	5.157	3.514
Wife has no education (yes=1)	524	0.256	0.437
Wife has some primary (yes=1)	524	0.273	0.446
Wife has completed primary (yes=1)	524	0.397	0.490
Wife has at least some secondary (yes=1)	524	0.074	0.263
Wife's years of education	524	3.435	3.229
Age of Husband	478	37.464	10.110
Age of Wife	524	32.648	8.843
Household size	524	5.708	2.123
# of kids < 6 years	524	1.403	0.958
# of kids aged 6-12 years	524	1.187	1.031
# individuals aged > 12 years	524	3.115	1.347
Household owns dwelling (yes=1)	524	0.937	0.243
Household owns land (yes=1)	524	0.840	0.367
Household has good floor (yes=1)	524	0.099	0.299
Household has good roof (yes=1)	524	0.210	0.408
# of sleeping rooms	524	2.076	1.017
Household has access to piped water (yes=1)	524	0.078	0.269
Household has improved latrine (yes=1)	524	0.073	0.260

Notes to Table: The table includes households resident in the same village over both rounds of the IFS-Mai Mwana survey, and where the main respondent was married, and either the head or spouse of her household. Data for some husbands is missing if they are not living in the household at the time of the survey, but are still married to the wife.

are in general poor, as indicated by their poor quality housing, and extremely limited access to water and sewerage infrastructure.

3.2 Defining the Risk Sharing Group

Having described the data, we now discuss how we define the potential risk sharing group. As noted above, formal financial markets are almost absent in Mchinji, and there was no government safety net in place at the time of the surveys.[25] Instead, existing research in anthropology and sociology indicates that social connections, and in particular, extended family connections play a critical role in helping households deal with the consequences of risk and adverse events: for example, Trinitapoli et al. (2014) documents the role of older siblings in protecting educational investments of younger siblings, while Peters et al. (2008) and Munthali (2002) document the essential role

---

[25] A cash transfer program, the Mchinji Cash Transfer, was being piloted in a small number of villages in Mchinji at the time of the survey. Less than 3% of households in our sample report receiving the transfer.

Table 3: Number of potential sources of support following adverse idiosyncratic event

Source of Support	Mean	Median	Std. Dev
Family	1.69	2	1.68
Friends	1.94	1	2.31
N	1048		

Notes to Table: This table shows the number of different individuals with a specific social relationship that a household expects to receive help from if it experiences an income loss as a result of an idiosyncratic adverse event.

Table 4: Transfers Given to and Received From Family and Friends

Source of Support	Support Given		Support Received		Support Given + Received	
	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev
Family	375.11	1485.83	321.22	1567.91	696.78	2378.13
Friends	113.59	677.72	87.65	599.74	201.24	919.48
N	1048		1048		1048	

Notes to Table: This table shows the amounts given to (left panel), received from (middle panel), and given to and received from (right panel) individuals with a specific social relationship by the household in the year prior to the survey for wave 1 and between surveys for wave 2. All amounts are in Malawi Kwacha. The exchange rate at the time of the survey was around US$1 = 140 MK.

played by the family in fostering and taking care of children orphaned by HIV/AIDS. We also find support for this in our data. In particular, looking at responses to a question on who households expect to receive informal monetary transfers, loans or gifts from, in the event of an income loss due to adverse idiosyncratic events (displayed in Table 3), we see that at the median, households expect to receive support from 2 family members and 1 friend. The average indicates the opposite pattern, though this is driven by a small number of households who can turn to a large number of friends.[26]

Our data also allow us to look at the actual amounts of transfers, loans or gifts (monetary or in-kind) given to and received from family and friends (displayed in Table 4) in the year prior to the survey. The data indicates, on average, households give around 375 MK to family, and receive on average 321 MK. Their transactions with friends are of a much lower magnitude (two and a half times, in fact), with 113 MK given on average and 87 MK received from friends. These pieces of evidence thus confirm that the extended family is a critical source of risk sharing in this setting. Given the importance of family for risk sharing in this setting, we define 'potential group size' based on family.

Further anthropological evidence allows us to define the potential group more finely, and also suggests a placebo test to rule out any potential lingering endogeneity concerns related to this definition. Within the family, anthropological evidence suggests that a wife's brothers should play an important role in ensuring the well-being of her family. The predominant ethnic group in our sample, the Chewa, are a matrilineal and matrilocal ethnic group (Richards, 1950, Phiri, 1983, Mtika and Doctor, 2002). Traditionally, under matriliny, society gives a special role to an individual's

---

[26] 1% of households report being able to turn to 10 or more friends in case of an adverse event.

maternal family, resulting in a close bond between siblings, even after marriage. Moreover, a woman's brothers play a crucial role in supporting her family: The eldest brother is responsible for ensuring access for a woman's family to production resources, healthcare, and other things important for household welfare. As a result, children will consult with their maternal uncles as they are responsible for arranging marriages, ensuring the children have access to adequate land and other productive resources, as well as health care (Phiri, 1983,Mtika and Doctor, 2002).

The literature indicates that some practices may be less relevant today, while other aspects of matriliny have proved to be remarkably resilient over time. For instance, the practice of matrilocality – whereby the husband moves to the wife's home immediately after marriage – has waned somewhat in Mchinji, with about a half of couples in our sample living in the husband's village when interviewed, and the other half live in the wife's village of birth. At the same time, though, children are still considered to 'belong' to their mother's matriline, and the maternal relatives become their key caretakers following her death Munthali (2002).

In terms of risk sharing arrangements, data on interhousehold transfers from the Family Transfers Project (collected within the Malawi Longitudinal Study of Families and Health) indicates that a wife's brothers remain an important source and recipient of transfers from a household: 33% (41%) of couples report having received (given) a material transfer from (to) the wife's brothers in the past growing season (which corresponds to a period of around 3-5 months). Moreover, they are less likely to receive material transfers from a wife's sisters (26% report receiving a material transfer), and received transfers are of lower magnitude (351 MK on average is received from brothers, compared to 119 MK from sisters).[27] The evidence thus suggests that the brothers of the wife are likely to still play an important role in risk sharing for the household. We thus define the potential risk sharing group to be the number of brothers (and separately, sisters) of the husband and wife.

3.3 Crop Losses

3.3.1 Measuring Crop Losses

Unexpected crop losses are used as our measure of shocks in the analysis.[28] Such crop losses could occur as a result of pests, variation in weather (whose effects could vary within a village by the type of soil, and other characteristics of the land), and other such factors. The first (second) survey collected information on whether the household experienced any crop loss in the year preceding the survey (or since the first survey); and if so, how much potential revenue was lost.[29] We use this information to construct two measures of crop loss: the first is a dummy variable defined to be 1 if the household experienced a crop loss event, thereby measuring the incidence of a crop loss; while

---

[27] These figures come from 220 observations, and are not adjusted for the number of siblings, or other variables.

[28] Crop losses have been used as a measure of adverse events by studies including Beegle et al. (2006).

[29] The exact questions were as follows: *"In the last year (since the last survey) did this household suffer from a bad harvest or crop loss?"* and *"How much potential revenue was lost as a result of the loss?"*

Table 5: Crop Losses, By Year

	N	Mean	Std Dev
Overall Sample			
Crop loss incidence	1048	0.242	0.429
Income lost ('000s MK)	1044	3.756	19.337
2008-09			
Crop loss incidence	524	0.303	0.460
Income lost ('000s MK)	524	5.536	26.310
2009-10			
Crop loss incidence	524	0.181	0.386
Income lost ('000s MK)	520	1.962	6.891

Notes to Table: Sample includes households resident in the same village across the two surveys, and where the main respondent was married at the time of the survey and either the head or spouse of the head.

the second is potential revenue lost normalised by a measure of 'permanent' consumption, thereby capturing the intensity of the crop loss.[30]

Crop losses are prevalent in this setting, as can be seen from Table 5: Around 24% of households in our sample experienced a crop loss over the 2-year period, losing on average, just over 3,700 MK. This amount corresponds to around one third of average monthly household food consumption. Among those who experienced a loss, the average loss is around 13,000 MK, which corresponds to 125% of average monthly household consumption. More crop losses were observed in the year prior to the 2008-09 survey relative to 2009-10, with the losses experienced in the former year being more severe in intensity.

Finally, there is some persistence in crop losses among those who experienced a loss. From Table 6, we see that around 8% of households experience a crop loss in both survey rounds, which is higher than what we would expect if crop losses were independently distributed.[31]

---

[30] We normalise the potential revenue lost by the household's permanent consumption to account for the fact that households that experience larger losses may be wealthier and better able to build up buffer stocks to deal with the consequences of risk. In this case, we would erroneously conclude that households are well insured. Household permanent consumption is measured as the part of household consumption predicted by the education of the female main respondent as measured in 2004. We also experimented with using household asset holdings in 2004 and quality of house in 2004, in addition to the education of the female main respondent, to predict household consumption. A concern with using past household assets, however, is that they may be correlated with a household's ability to currently smooth consumption, particularly if crop loss events are persistent. Results using this measure are available on request.

[31] Under the assumption that the crop loss distributions for the two years are independent, the probability of experiencing a crop loss in both survey rounds is the product of the probability of experiencing a crop loss in 2008-09 and the probability of experiencing a crop loss in 2009-10, which equates to around 5.4% of households.

Table 6: Persistence of Crop Losses

		Crop Loss in 2009-10		
		No	Yes	Total
Crop Loss in 2008-09	No	312	53	365
		[59.54]	[10.11]	[69.66]
	Yes	117	42	159
		[22.33]	[8.02]	[30.34]
		429	95	524
		[81.87]	[18.13]	[100]

Notes to Table: Sample includes households resident in the same village across the two surveys, and where the main respondent was married at the time of the survey and is either the head or spouse of the head. Percentages in each category displayed in the parentheses.

3.3.2 Are crop losses idiosyncratic within the village?

Our objective is to investigate how the amount of idiosyncratic risk shared by a household varies with the size of its extended family. For our tests to have sufficient power, we require that there is sufficient variation within villages in the incidence of crop losses.[32] Such variation may arise as a result of differences in land quality, with some plots more resilient to poor weather relative to others; or due to variation in the crops grown (some crops and crop varieties may be more resilient to poor weather); or due to localised pests or crop diseases. Note that there was no drought or widespread flooding in Mchinji over the survey period. Nonetheless, we check here for the amount of idiosyncratic variation in our data. To do this, Figure 3 displays histograms of the within-village variation in the incidence of a crop loss, for each round of data. We see from the Figure that there are a number of villages with idiosyncratic variation in the incidence of crop losses.

3.4 Measuring extended family networks

To investigate the relationship between the extent of risk sharing and the size of the extended family, we collected information in the survey on the numbers of siblings of the main respondent and her spouse. Data were collected on the numbers of siblings in the village and the number living[33], and on the location of residence of the respondent's mother and mother-in-law. We use the numbers of siblings as our measure of potential group size. The two surveys – conducted around a year

---

[32] As we will show below, ideally we would like to be able to control for within-group shocks. However, we are unable to do this since we do not observe information on all members of the group. Controlling for aggregate village shocks allows us to partially account for common shocks experienced by group members in the village.

[33] The exact wording of the questions was as follows: Please tell me how many of the following categories of relatives are currently alive, regardless of where they live:

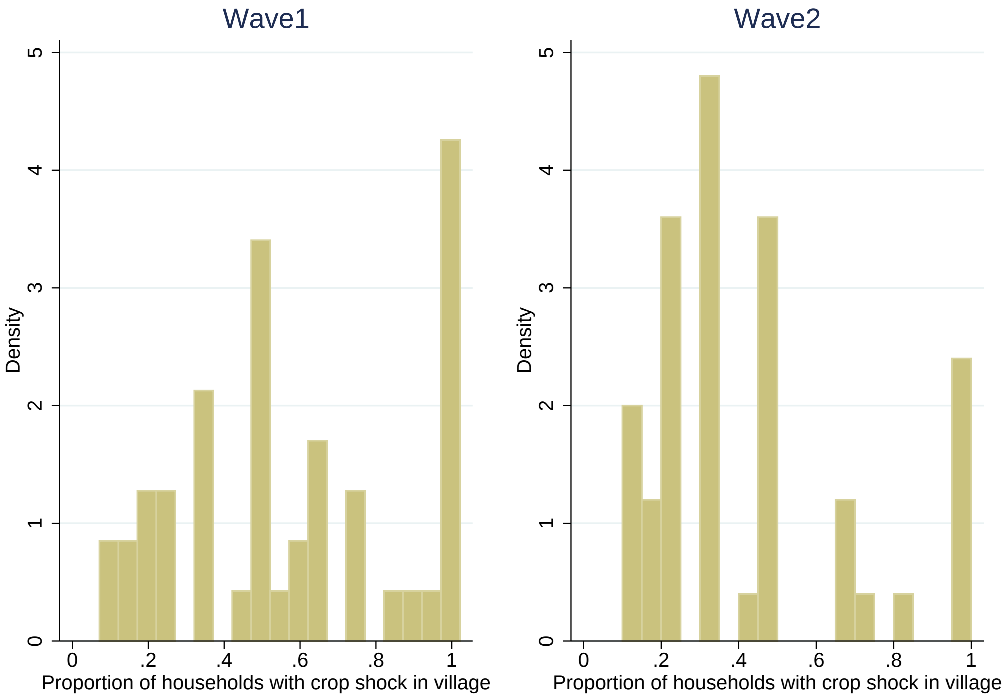
1. Sisters 2. Sisters-in-law 3. Brothers 4. Brothers-in-law

Please tell me how many of the following categories of relatives are currently living in this village:

1. Sisters 2. Sisters-in-law 3. Brothers 4. Brothers-in-law

Note that in our survey, sisters-in-law and brothers-in-law were translated in a manner so as to capture the siblings of one's spouse.

Figure 3: Variation in crop loss incidence within villages



Notes to Figure: The Figure plots a histogram for the proportion of households in each village that experienced a crop loss in wave 1 of the survey (left panel) and in wave 2 (right panel). For legibility of the graph, a peak at 0 with magnitude 10 has been omitted.

Table 7: Any Family Links

	Any Sibling Link	Any Sibling Link of Husband	Any Sibling Link of Wife	Any Links			
				Husband		Wife	
				Brothers	Sisters	Brothers	Sisters
Alive	0.996	0.971	0.985	0.908	0.908	0.941	0.933
	[0.003]	[0.008]	[0.005]	[0.013]	[0.013]	[0.011]	[0.012]
In Same Village	0.819	0.666	0.531	0.534	0.517	0.418	0.437
	[0.021]	[0.024]	[0.028]	[0.025]	[0.022]	[0.021]	[0.030]

Notes to Table: The table includes households resident in the same village over both survey rounds, and where the main respondent is married, and is either the head or spouse of her household.

apart – captured similar numbers of siblings for a large part of our sample. However, there were some discrepancies in a sizeable minority (˜30%) of observations, which could not be explained by naturally expected changes (e.g. deaths or divorce), and thus point towards reporting errors. To mitigate effects of such errors, we take the average of the reported information in both surveys as the preferred measure of potential group size. Moreover, we use information from the household roster, along with this data to construct variables for the number of siblings of the husband and wife living outside the household.

Tables 7 and 8 provide some descriptive statistics of sibling networks in this context. Virtually all households have a sibling link outside the household, and a lower, though sizeable proportion (˜82%), has siblings within the same village. Households have on average 9.4 siblings outside the household, of whom close to 3 are within the same village. The high numbers of siblings (relative to Western contexts) reflects the high fertility rates in Malawi: the Total Fertility Rate[34] in rural areas was estimated to be around 7.6 in 1984, falling slightly to 6.7 by 2000. At the individual level, almost all husbands and wives have a living sibling, though roughly one-third of husbands and nearly half of wives do not have a sibling in the same village. On average, wives have more living siblings (˜5) than husbands (˜4.4), but both have similar numbers of siblings in the same village.

These patterns are in line with post-marital living patterns in this context. As mentioned already, though the Chewa were traditionally matrilocal, this seems to be waning in Mchinji, with roughly half of the wives in our sample moving to their husbands' village after marriage. Thus, roughly half the wives in our sample have a sibling in the same village, while two-thirds of husbands have a sibling in the same village. In terms of the type of sibling link, husbands and wives have similar numbers of brothers and sisters alive, though they have slightly more brothers than sisters living in the same village.

---

[34] This captures the average number of children that would be born to a woman over her lifetime if she were to experience the exact current age-specific fertility rate through her lifetime, and if she were to survive from birth to the end of her reproductive life.

Table 8: Numbers of Family Links

	# of Sibs of Husband + Wife	# of Sibs of Husband	# of Sibs of Wife	Number of Husband Brothers	Number of Husband Sisters	Number of Wife Brothers	Number of Wife Sisters
Alive	9.418	4.422	5.162	2.281	2.267	2.519	2.740
	[0.172]	[0.098]	[0.113]	[0.064]	[0.068]	[0.069]	[0.079]
In Same Village	2.945	1.571	1.498	0.893	0.788	0.811	0.748
	[0.127]	[0.081]	[0.086]	[0.057]	[0.044]	[0.050]	[0.052]

Notes to Table: The table includes households resident in the same village over both survey rounds, and where the main respondent was married, and either the head or spouse of her household.

4 Empirical Model

Our objective is to understand how the amount of risk shared in the face of crop losses varies with the size of a household's family network. To do so, we require a measure of risk sharing, which can be computed in the available data. One measure implied by the model (assuming utility of the constant relative risk aversion form) is the deviation of changes in log consumption from the first-best allocation. Under the first-best allocation, where every group is stable, each household will consume an equal share of pooled resources. This means that changes in household-level log consumption should move one-to-one with aggregate group resources, and be uncorrelated with household-level idiosyncratic shocks. This is a well known result in the risk sharing literature (see, for example, Townsend, 1994), which we use to construct our test for how risk sharing varies with the size of a household's family network.

Using consumption to construct our measure of risk sharing has the advantage of providing a useful summary measure of all the different risk sharing strategies employed by a household. Collecting reliable information on all the different methods used for risk sharing, and of the exact bilateral transactions between households in a group is very time-consuming and costly; and more vulnerable to measurement error: For example, Mtika and Doctor (2002) report that one reason why households in Malawi report few transfers to their parents is that respondents help out their parents all the time and do not remember all of the details of specific transactions; while Comola and Fafchamps (2015) show that there is a strategic behaviour in reporting bilateral inter-household transfers in rural Tanzania.

We next describe our estimation equation. The theoretical model did not suggest any clear prediction on the shape of this relationship. We thus begin by estimating a non-parametric relationship between group size and the extent of risk sharing. We do so using the following equation, which includes interaction terms with dummy variables for each potential group size value in the data:

$$\Delta log(c_{ivt}) = \alpha_0 + \alpha_1 \Delta(crop_{ivt}) + \sum_{n=1}^{N} \beta_n \Delta crop_{ivt} * 1(S_{iv} = n) + \Delta X_{ivt}\gamma$$
$$+ \sum_{n=1}^{N} \lambda_n \Delta crop_{ivt} * 1(F_{iv} = n) + \mu_{vt} + \Delta\epsilon_{int} \quad (6)$$

where $\Delta log(c_{ivt})$ is the change over time in log consumption for household i in village v at time t, $\Delta(crop_{ivt})$ indicates the change in crop loss incidence or intensity for household i between t and $t-1$, where the crop loss incidence and intensity are measured as explained in Section 3.3. The term $1(S_{ivt} = n)$ takes the value of 1 if the household has n brothers or sisters of the head or spouse and 0 otherwise. ΔX_{ivt} captures changes in household characteristics, such as household demographics, that could also affect changes in log consumption. The term $\sum_{n=1}^{N} \lambda_n \Delta crop_{ivt} * 1(F_{iv} = n)$ controls for direct effects of total sibship size of the husband or wife. μ_{vt} denote village-time dummies which capture village-level aggregate shocks. The coefficients of interest are β_n, while the sum of the coefficients $\alpha_1 + \sum_{n=1}^{N} \beta_n * 1(S_{iv} = n)$ indicates how well protected a household's consumption is against idiosyncratic crop losses. In line with the prevailing social norms in this context which indicate that a woman's brothers have an important role in helping out their sisters' households, we conduct the empirical analysis separately for the brothers and sisters of the head of a household and his spouse.

Ideally, we would like to control for group-level aggregate shocks, rather than just village-level aggregate shocks. However, we are unable to do so since we do not observe the crop losses or consumption of all members of the potential group. As a result, the group-level aggregate shock is an omitted variable, which will bias the estimates of interest if it is correlated with potential group size or crop loss incidence. To assess the consequences of this, we run some simulations where we generate data from a data generating process similar to that implied by the model in Section 2 (parameterised using values similar to those in the data), and use these to shed light on the direction and magnitude of the resulting omitted variable bias. The findings of this exercise are given in Subsection 5.2.

We include changes in crop loss, rather than crop loss in levels, as a measure of idiosyncratic shock for the following reason: assume we used the crop loss incidence between periods t and $t+1$ as the shock measure. The concern with this is that, in the absence of perfect risk sharing, a household may already have low consumption at period t if it experienced a crop loss between periods $t-1$ and t. Moreover, assume it experiences another crop loss between t and $t+1$, and its consumption remains low at time $t+1$, resulting in little or no change in $\Delta log(c_{hvt})$. The household would then erroneously appear to be perfectly insured: so if crop losses are persistent (and there is some evidence of this for some households as seen in Section 3.3), we would erroneously conclude that households are perfectly insured since their consumption does not respond to a crop loss. For this reason, we define the shock measure as the difference in incidence (or intensity) of a crop loss between time periods $t-1$ and t and between t and $t+1$.[35]

This specification can shed light on the shape of the relationship between our measure of risk sharing and the size of a household's potential group. However, this approach, which is fully non-parametric in the number of siblings, might not have sufficient power to identify statistically significant effects. To improve power, we divide potential group size into three bands, the cutoffs of which are motivated by the findings from the nonparametric regression above, and use the following specification for the empirical analysis:

$$\begin{aligned}\Delta log(c_{hvt}) = \alpha_0 + \alpha_1 \Delta(crop_{hvt}) + \sum_{g=1}^{G} \beta_n \Delta crop_{hvt} * 1(NS_{g,hv} = 1) + \Delta X_{hvt}\gamma \\ + \sum_{n=1}^{N} \lambda_n \Delta crop_{hvt} * 1(F_{hv} = n) + \mu_{vt} + \Delta\epsilon_{hnt} \qquad (7)\end{aligned}$$

where $1(NS_{g,hv} = 1)$ is a term that takes value 1 if the household's network size is within the cutoffs associated with band g, and 0 otherwise; and the rest of the variables are as defined above.[36]

5 Results

5.1 Main Specification

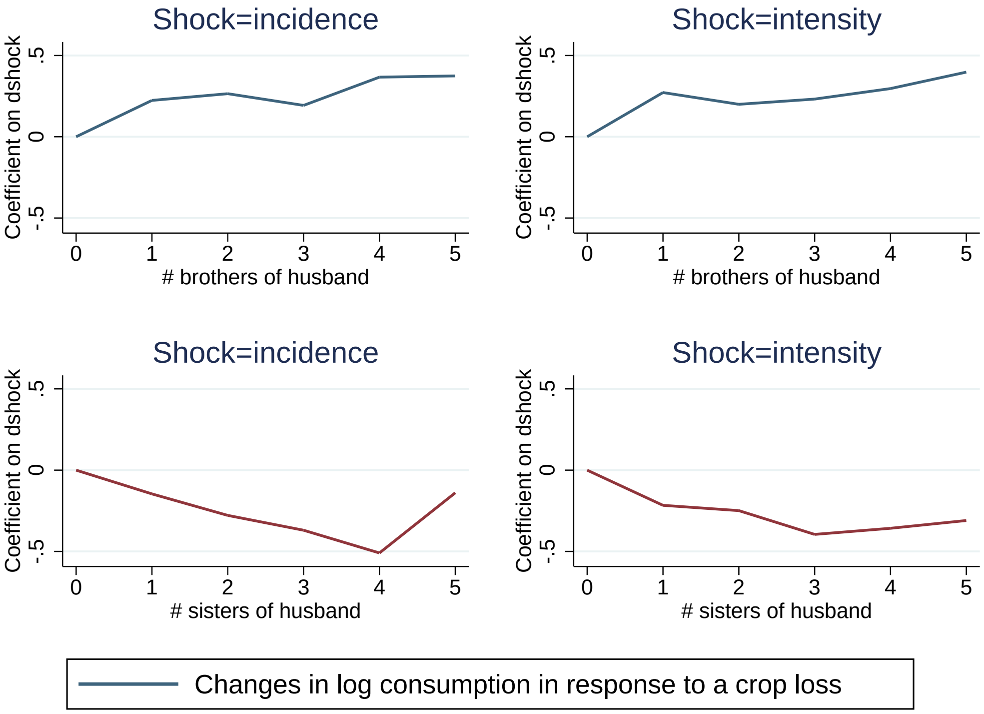
We first estimate Equation 6, separately for the brothers and sisters of the husband and wife. Figures 4 and 5 plot the coefficients from these regressions. We have extremely limited power in these specifications, and thus suppress the confidence intervals for these coefficients from the Figures. Despite the limitations in power, these Figures shed light on the possible shape of the relationship between informal risk sharing and potential group size in our data.

---

[35] A further issue with focusing on incidence of rather than changes in crop losses is that we do not account for other risk faced by the household, which may affect both their consumption smoothing and the shocks they experience. To assess the importance of this issue in our context, we estimated specifications controlling for other idiosyncratic shocks experienced by the household (business shocks, theft, and marriage break-up) and found it made little difference to the key coefficients of interest.

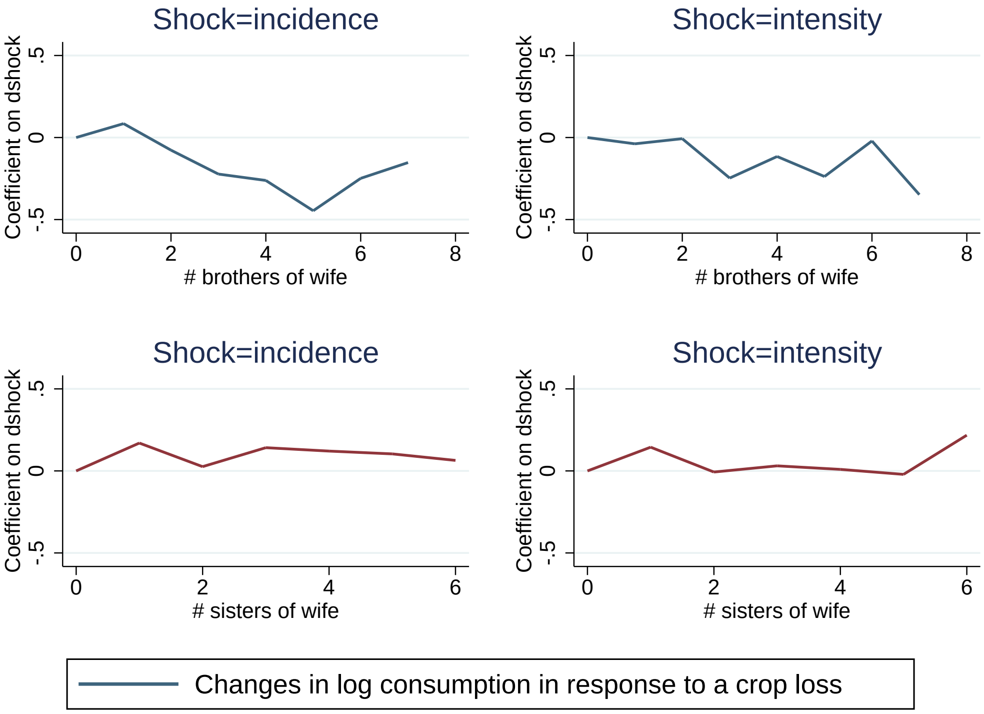
[36] The exact cutoffs for the different bands are defined in Section 5.

Figure 4: Risk Sharing by Number of Brothers and Sisters of Husband



Notes to Figure: The figures plot the correlation between changes in log consumption and household crop loss incidence (left panel) and intensity (right panel) for households with different numbers of brothers (top panel) and sisters (bottom panel) of the husband. The coefficient for zero brothers or sisters is normalised to 0, and lower values of the coefficient indicate worse risk sharing.

Figure 5: Risk Sharing by the Number of Brothers and Sisters of Wife



Notes to Figure: The figures plot the correlation between changes in log consumption and household crop loss incidence (left panel) and intensity (right panel) for households with different numbers of brothers (top panel) and sisters (bottom panel) of the wife. The coefficient for 0 brothers or sisters is normalised to 0, and lower values of the coefficient indicate worse risk sharing.

From the Figures, we can see that there are differences in the amount of consumption smoothing in the face of crop losses, by the size and types of family relations. In particular, Figure 4 indicates positive changes in log consumption (implying better protection of consumption) with larger numbers of brothers for the husband, and worse consumption smoothing with larger numbers of sisters of the husband. For the siblings of the wife, the Figures indicate that the consumption of households where the wife has a small number of brothers is almost perfectly smoothed, but worsens as the number of brothers increases. By contrast, no such relationship is seen for the number of sisters of the wife.

The analysis above suggests that there are nonlinearities in the relationship between the amount of risk shared in response to crop losses and the number of brothers and sisters of a household's head and spouse. Moreover, in line with the social norms suggested by the literature, the effects vary by type of sibling. In particular, for brothers of the wife, and sisters of the husband, there is an initial improvement in consumption smoothing with network size, before worsening. However, we do not have sufficient power to obtain statistically significant estimates. To gain power, we thus pool together the number of siblings of a particular type into 3 groups: those with 0 siblings of a particular type, those with 1-2 siblings of that type, and those with 3 or more siblings of that type.

These cutoff values are in line with the evidence presented in Figures 4 and 5 above, while also ensuring that each group has sufficient sample size to improve power. Table 9 presents the results for this specification, with our two measures for the crop loss shock: incidence and intensity. The top left panel displays the results pertaining to the brothers of the husband, while the top right panel displays these for the brothers of the wife. The bottom panel displays the results respectively for the sisters of the husband (left panel) and wife (right panel).

The regression coefficients indicate that households where the wife has more than 3 brothers experience much worse risk sharing following crop losses than those where she has fewer than 3 (i.e. 0 or 1-2) brothers. We detect no such relationship for the brothers of the husband. This finding is replicated across both our measures of crop losses - incidence and intensity. The coefficient estimates indicate that households where the wife has more than 3 brothers cut their consumption by approximately 26% when hit by a crop loss, while the intensity measure indicates that a crop loss of a magnitude equivalent to a month's consumption leads to a reduction in household consumption of approximately 18%.

The bottom panel of the table indicates worse risk sharing (significant only for the intensity measure) among households where the husband has any sisters, as can be evidenced by the positive coefficient on the interaction term for no sisters, and the negative coefficient on the interaction term for more than 3 sisters. No similar pattern is found for the sisters of the wife or the brothers of the husband. The absence of any significant differences in risk sharing by the number of sisters of the wife is consistent with the evidence showed in Subsection 3.2 which indicated that sisters of the wife are less important for risk sharing.

5.2 Robustness

The analysis in the previous subsection indicates that households where the wife (husband) has many brothers (sisters) achieve worse risk sharing following an idiosyncratic crop loss. In this Subsection, we outline various exercises undertaken to ascertain the robustness of this finding. In particular, we rule out that this finding is a result of being unable to account for unobserved common group shocks, or because larger networks are poorer, or because there is higher competition for resources amoung networks with many males, or because larger networks are more vulnerable to crop losses.

5.2.1 Aggregate Extended Family Shocks

As mentioned above, our data doesn't allow us to adequately account for common shocks at the extended family level. Such common shocks might be correlated with potential group size, hence biasing our estimates. For example, larger potential groups might be less vulnerable to a common group shock than smaller potential groups, leading to a negative bias in our coefficients of interest if the common group shock is not accounted for.

Table 9: Main results

	[1] $\Delta\ logc_{int}$	[2] $\Delta\ logc_{int}$	[3] $\Delta\ logc_{int}$	[4] $\Delta\ logc_{int}$
	Siblings of husband alive		**Siblings of wife alive**	
	shock = crop	**shock = Loss/Pred. Cons**	**shock = crop**	**shock = Loss/Pred. Cons**
Δshock	0.2114**	0.1915*	0.0126	0.0216
	[0.1003]	[0.0972]	[0.1294]	[0.2552]
No brothers*Δshock	-0.2306	-0.184	0.0264	0.0296
	[0.1690]	[0.1323]	[0.1482]	[0.0792]
≥ 3 brothers*Δshock	0.0015	0.0614	-0.2578**	-0.1786***
	[0.1029]	[0.0492]	[0.1129]	[0.0615]
N	524	519	524	519
R-squared	0.3213	0.3348	0.3216	0.3366
Δshock	-0.1609	-0.2151	0.1095	0.1061*
	[0.1809]	[0.1403]	[0.1453]	[0.0618]
No sisters*Δshock	0.1594	0.2218**	-0.0783	0.0571
	[0.1231]	[0.0997]	[0.1246]	[0.3093]
≥ 3 sisters*Δshock	-0.1522	-0.1274**	0.042	-0.0411
	[0.1039]	[0.0515]	[0.1068]	[0.0546]
N	524	519	524	519
R-squared	0.3126	0.3288	0.3253	0.3371

Notes: *** Significant at the 1% level; ** the 5% level; * the 10% level. Standard errors clustered at the village level in parentheses. Regressions pool together all households where a married head or spouse was surveyed, and who were resident in the same village for both survey rounds. All specifications control for village-time dummies and changes in household demographics. “Crop” indicates whether or not a household suffered a crop loss, while “Loss/Pred. Cons” measures the intensity of the crop loss as the income lost normalised by predicted household consumption.

We use simulations to assess the magnitude and sign of this bias, under different assumptions on the magnitude of the common extended family level shock. We generate data under the assumption that risk is shared according to the model in Section 2 (augmented to allow for group-specific shocks), and parameters are set to match those in our data (where possible). In particular, we set the group size distribution to be match the empirical distribution of brothers of the wife. Household income, y_i, consists of two components: y_h, an idiosyncratic household level component, and y_g - a common extended family component. Households' consumption rules are estimated numerically from the model in Section 2 augmented to allow for an independent group-level shock. For different levels of the common family shock, and randomly drawn idiosyncratic and group shocks, we assess how the coefficients on a specification similar to Equation 7 change when we add controls for common extended family shocks, rather than for common village shocks.[37] The findings are displayed in Table 10. The table indicates that all the coefficients are indeed biased, as expected. Moreover, the biases are sizeable, ranging from 10% of the true value to over 200%. In terms of the sign of the bias, α_1, the coefficient on the $\Delta crop$ variable is biased downwards, while β_2 (coefficient on $\Delta crop_{ivt} * 1(NS_{g,iv} = 1)$) is biased upwards in all but one case. By contrast, the coefficient on $\Delta crop_{ivt} * 1(NS_{g,iv} \geq 3)$,$\beta_3$, is biased upward. So, if anything, we are likely to be *underestimating* the negative effect of larger groups on risk sharing.

5.2.2 Are larger networks poorer?

An important concern is that our findings may be driven by unobserved factors that drive both network size and changes in log consumption. One such set of factors relates to the fact that households with larger family networks may be poorer. Larger families have long been observed to be poorer in a variety of contexts. This could make them less able to provide support to other family members when they need it, thus leading to worse risk sharing. We provide evidence that our results are not driven by the fact that larger families are poorer.

First, we fail to find similar results for the sisters of the wife, and for the brothers of the husband. If the findings were being driven by a family size effect, rather than being the effect of having many brothers, we would expect to find that households with many sisters are also less well protected from crop loss events. Of course, this argument is only valid as long as households with many sisters and those with many brothers are not different in other dimensions. To assess whether this is the case, we test whether households where the wife has ≥ 3 brothers and < 3 sisters are different to households with ≥ 3 sisters, but < 3 brothers, focusing on dimensions that are less likely to have changed as a result of recent shocks experienced by households. The findings from this analysis are displayed in Table 11 (12) for the husband (wife).

As can be seen from the tables, we find few differences in the small set of observable characteristics of the husband and wife in these two types of households. In particular, there are no significant differences in the amount of education of the husband or wife in households where the husband (wife)

---

[37] Full details on the simulations and estimation equation are given in Appendix B.

Table 10: Simulation Results to Assess the Sign and Magnitude of the Bias from Omitting Controls for Aggregate Extended Family Shocks

	Size of Group-Level Shock (% of Avg. Annual Household Income)					
	0%	5%	10%	15%	20%	25%
Avg. $\hat{\alpha_1}$ (Group dummies)	-0.046	-0.183	-0.105	-0.105	-0.105	-0.117
Simulation std. error	[0.001]	[0.003]	[0.001]	[0.001]	[0.001]	[0.001]
Avg. $\tilde{\alpha_1}$ (Village dummies)	-0.119	-0.213	-0.160	-0.161	-0.161	-0.169
Simulation std. error	[0.003]	[0.003]	[0.003]	[0.001]	[0.003]	[0.004]
Avg. % Absolute Bias	160.7%	16.0%	51.6%	52.3%	52.4%	44.4%
Avg. $\hat{\beta_1}$ (Group dummies)	-0.106	0.090	-0.031	-0.086	-0.085	-0.099
Simulation std. error	[0.002]	[0.004]	[0.001]	[0.001]	[0.001]	[0.002]
Avg. $\tilde{\beta_1}$ (Village dummies)	-0.092	0.032	-0.044	-0.072	-0.072	-0.078
Simulation std. error	[0.004]	[0.004]	[0.004]	[0.004]	[0.004]	[0.006]
Avg % Absolute bias	13.4%	64.9%	42.9%	15.8%	15.8%	21.1%
Avg. $\hat{\beta_2}$ (Group dummies)	-0.074	0.037	-0.051	0.009	0.010	0.009
Simulation std. error	[0.002]	[0.004]	[0.002]	[0.003]	[0.002]	[0.003]
Avg. $\tilde{\beta_2}$ (Village dummies)	-0.033	0.040	-0.021	0.029	0.029	0.028
Simulation std. error	[0.003]	[0.003]	[0.003]	[0.004]	[0.004]	[0.005]
Avg % Absolute bias	56.2%	9.5%	58.6%	225%	217.5%	220%

Notes to Table: Data simulated with parameters to match those in data. Exact parameter values, and simulation details are explained in Appendix B. The average annual household income is around 56000 MK. $h_h = 61223$MK and $l_h = 46475.64$MK. α_1 is the coefficient associated with $\Delta crop$, β_1 is that associated with No sibling of that type * $\Delta crop$; and β_2 is that associated with ($\geq$3 siblings)*$\Delta crop$.

Table 11: Comparing characteristics of households where husband has $\geq$ 3 brothers with those where he has $\geq$ 3 sisters

	≥3 sis of husband	**sd**	**≥3 bros of husband**	**sd**	**p-val of diff**
Husband's Characteristics					
Years of education	4.815	0.380	5.257	0.329	0.391
Age	37.865	0.923	37.269	0.814	0.632
Chewa	0.931	0.027	0.945	0.028	0.527
Wife's Characteristics					
Years of education	3.404	0.282	3.609	0.274	0.582
Age	33.685	0.839	33.027	0.663	0.525
Chewa	0.978	0.014	0.973	0.014	0.768

Notes: ** Significant at 5% level; * at the 10% level. Sample includes households where the wife has 3 or more brothers and less than 3 sisters or 3 or more sisters and less than 3 brothers.

Table 12: Characteristics of households where wife has $\geq$ 3 brothers with those where she has $\geq$ 3 sisters

	≥3 sis of wife	**sd**	**≥3 bros of wife**	**sd**	**p-value of diff**
Husband's Characteristics					
Years of education	5.202	0.322	5.519	0.377	0.517
Age	37.487	0.807	37.404	1.051	0.954
Chewa	0.915	0.035	0.886	0.052	0.392
Wife's Characteristics					
Years of education	3.760	0.277	3.872	0.326	0.794
Age	33.241	0.592	32.456	0.925	0.489
Chewa	0.962	0.017	0.972	0.016	0.352

Notes: ** Significant at 5% level; * at the 10% level. Sample includes households where the husband has 3 or more brothers and less than 3 sisters or 3 or more sisters and less than 3 brothers.

has $\geq$ 3 brothers and those where he (she) has $\geq$ 3 sisters. Though males typically have a higher level of education than females, there is no difference in education levels by the sex composition of the individual's sibship.[38]

5.2.3 Number of Brothers and Competition for Resources

Another concern is that there might be more competition for production resources among families with many males: essentially, if land is passed down to males only, and there are many males in a particular family, each male would receive a smaller land plot, and thus would be less able to help their sisters' households when they face idiosyncratic shocks. The land descent system in Mchinji is considered to be a mixed one: some households practice a patrilineal system and pass on land to males, whereas others practice a matrilineal system and pass on land to females. We provide some suggestive evidence to rule out this channel. In particular, though we do not have information on the landholdings of siblings of the husband or wife, we can look at whether there are any differences in the size of land between households where the husband has many brothers and few sisters compared to those where the husband has many sisters but few brothers. If the patrilineal form of land descent is more dominant in our sample (which we do not believe it to be), we would expect households where husbands have many brothers to have smaller plots of land than households where the husband has many sisters. Examining the data, we see that households where the husband has 3 or more brothers and fewer than 3 sisters have on average 2.9 hectares of land, whereas those where the husband has 3 or more sisters and fewer than 3 brothers have on

---

[38] The differences in education levels by gender are likely to be driven by gender differences in the economic returns to education rather than due to explicit gender discrimination by parents. To our knowledge, there is no evidence of sex discrimination in investments in children at either the pre-natal or post-natal stage. Indeed, when we analyse the effects of a randomised infant feeding counselling intervention in this context by gender, we find no differences in nutritional investments in children by gender. These results are available on request.

average 2.7 hectares of land. This difference is not statistically significant, thus providing suggestive evidence that the empirical findings are unlikely to be driven by this channel.

5.2.4 Potential Group Size and Incidence of Shocks

A second concern is that larger extended families could be more vulnerable to crop loss events, particularly if they are poorer. In that case, the deficiencies in risk sharing detected above may be a consequence of poverty, rather than a breakdown of risk sharing due to unstable coalitions.

To see if this is the case, we consider how the incidence and intensity of crop losses vary with potential group size. To do so, we regress the crop loss and intensity variables on our network size variables, pooling data from both survey rounds. Table 13 displays these results. The table does not indicate that households where the wife has many brothers are more vulnerable to crop loss events compared to households where the wife has fewer brothers. Thus, we can rule out that our finding of poor risk sharing among these households is driven by this channel. Interestingly, we find a negative coefficient for households where the wife has 3 or more sisters: such households are less likely to be affected by a crop loss incident, though there is no difference detected in the intensity of the crop loss.

Table 13: Network size and crop loss incidence

	[1] **crop loss incidence**	[2] **crop loss intensity**	[3] **crop loss incidence**	[4] **crop loss intensity**
	Siblings of husband alive		**Siblings of wife alive**	
No brothers	-0.0571	-0.0752	-0.0058	0.0012
	[0.0452]	[0.0721]	[0.0471]	[0.0902]
≥3 brothers	-0.014	-0.0527	0.0033	-0.0599
	[0.0262]	[0.0424]	[0.0275]	[0.0428]
N	1131	1131	1131	1131
R-squared	0.0244	0.0200	0.0289	0.0216
No sisters	0.0036	-0.0083	-0.0290	-0.0633
	[0.0548]	[0.0731]	[0.0522]	[0.0818]
≥3 sisters	-0.0075	-0.0203	-0.0628**	-0.0216
	[0.0285]	[0.0384]	[0.0314]	[0.0391]
N	1131	1131	1131	1131
R-squared	0.0262	0.0198	0.0306	0.0191

Notes: *** Significant at the 1% level; ** the 5% level; * the 10% level. Standard errors clustered at the village level in parentheses. Regressions pool together all households where a married head or spouse was surveyed and who were resident in the same village for both survey rounds. ”Crop loss incidence” is a dummy variable that indicates whether the household experienced a crop loss in the previous year (or since the last survey), while ”Crop loss intensity” is the size of the crop loss normalised by predicted household consumption.

6 Calibration

The empirical results show that households where the wife (husband) has a large number of brothers (sisters) achieve worse risk sharing outcomes compared to households where the wife (husband) has fewer brothers (sisters). The theory indicates that the relationship between risk sharing and potential group size is ambiguous and sensitive to parameter values: for some combination of parameters, larger potential groups can offer better risk sharing, while for others, they offer worse risk sharing. To investigate whether the findings can be explained by the theory, we conduct a calibration exercise to see if the model can reproduce the empirical findings when parameter values are set to be similar to those in our data.

We parameterise the value of the high and low endowment as follows: From the data, we obtain the average annual household income from agriculture for all households in the sample, $\bar{y}$. This is equivalent to a weighted average of the high and low endowment states, where the low endowment state is taken to be the high endowment state less the crop loss (in nominal terms, without normalising for predicted consumption):

$$\bar{y} = p * h + (1 - p)(h - crop) \tag{8}$$

We obtain the values for $\bar{y}$, p and $crop$ from the data, and use the formula 8 to back out the values for h and l respectively. Table 14 displays the resulting parameters. In addition to these parameters, we also need to specify a value for the coefficient of relative risk aversion, ρ and the discount factor, δ. We set $\rho = 1.5$ and $\delta = 0.95$. The value for δ, which is lower than that typically estimated for developed countries, is within the range estimated for India by Ligon et al. (2003).

Table 14: Parameter values for calibration

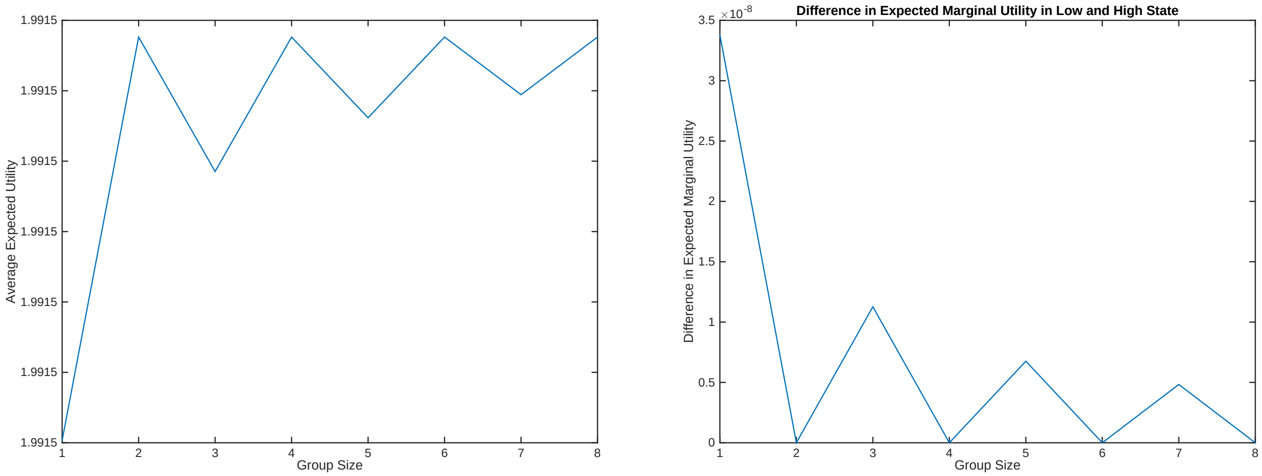
Parameter	Value
h	61223.64MK
l	46475.64MK
p	0.63
δ	0.95
ρ	1.5

Note to Table: This table displays the parameter values used to calibrate the theoretical model. The values for the high and low endowments, h and l are in Malawi Kwacha (MK). The exchange rate at the time of the survey was roughly US$1 = 140MK.

Figure 6 plots the value for average expected utility and group size. What is striking is that weighted expected utility increases with potential group size initially, but then falls before increasing again in a zigzag pattern. This pattern can be explained by the fact that given the parameter values, only groups of size 1 and 2 are stable. Larger potential groups would then sort randomly into the smaller stable groups, for example, groups of size 3 would sort into groups of size 1 and 2. Since expected utility under autarky is lower than in a group of size 2, this results in a drop in average expected utility for a potential group of size 3. In fact, such an argument holds for all odd-sized potential groups, while even-sized potential groups would sort into subgroups of 2 and attain the same average expected utility as a group of size 2.

Importantly, the drop in expected utility when moving from a potential group of size 2 to 3 matches the pattern found in the data, suggesting that threats of coalitional deviations may be a possible explanation for the worse risk sharing for households where the wife has many brothers.

Figure 6: Calibration Findings - Average Expected Utility and Network Size



Notes: The Figure on the left panel shows the relationship between weighted average expected utility and group size, while that on the right panel shows the relationship between the weighted average expected difference in marginal utility and group size

7 Conclusion

In this paper, we study the relationship between group size and the extent of risk sharing in a setting with limited commitment and coalitional deviations. In such environments, two forces are at play in determining the relationship between group size and risk sharing: on the one hand, larger groups allow for more effective diversification, and hence better risk sharing. On the other hand, they are more vulnerable to deviations by sub-groups (coalitions) of households who can renege on the informal arrangement and continue sharing risk in the smaller subgroup. Thus, risk sharing groups will be bounded from the top (GR). In this paper, we extend the model of GR and use simulations to show that the relationship between risk sharing and group size is theoretically ambiguous. The nature of this relationship is thus an empirical question.

We investigate this question empirically using data from rural Malawi, and overcome the challenge posed by the fact that the size of the actual risk sharing group is endogenous, by considering potential group size, and focusing on a group likely to be exogenous – siblings of the household head and spouse. Evidence from the anthropological and sociological literatures indicate that the extended family is a crucial risk sharing institution in the setting we study. Moreover, historical, well-documented norms at play in this context indicate a much more important role for a wife's brothers in providing risk sharing, than for her sisters. These norms highlight an important source of heterogeneity in risk sharing patterns, and also allow us to construct placebo tests to alleviate concerns that unobserved factors correlated with our measures of group size and the efficiency of risk sharing are driving our findings.

We consider how well protected a household's consumption is to idiosyncratic crop losses – an important source of risk in this setting – allowing the effects to vary by the size of the family network of the husband and wife (defined separately by gender of sibling). In line with the literature on informal risk sharing, we measure the degree of risk sharing by the correlation between changes in household log consumption and idiosyncratic crop losses (net of aggregate shocks at the village level). A first non-parametric specification, which places no restrictions on the shape of the relationship between the degree of risk sharing and potential group size, indicates that this relationship is non-linear. However, these estimates are extremely imprecise.

To increase power, we divide group size into three bands, the boundaries of which are informed by the non-parametric analysis. Estimates from this specification indicate that households where the wife (husband) has many brothers (sisters) achieve worse risk sharing relative to households where they have fewer brothers (sisters). We fail to find a similar relationship for the wife's sisters (brothers), which indicates that the relationship is unlikely to be driven by the fact that households where the husband/wife have many siblings are poorer. Moreover, we show that these households are not more susceptibe to crop losses, suggesting that the findings are not driven by this margin either. We also provide suggestive evidence to rule out other channels including higher competition for production resources among extended families with many male siblings. A calibration exercise, where we parameterise our theoretical framework using information from the data (where available), indicates that the empirical patterns could be produced by the theory.

Thus, larger potential risk sharing groups need not yield better risk sharing outcomes, indicating a role for governments and other actors to implement policies and mechanisms to better protect household wellbeing.

References

"Malawi Population and Housing Census". National Statistics Office, 2008.

Finscope malawi 2008 report. http://www.finance.gov.mw/fspu/index.php/studies-reports-fspu/completed-studies/52-2008-finscope-malawi-survey/file, 2009.

A. Ambrus, M. Mobius, and A. Sziedl. "Consumption Risk Sharing in Social Networks". *American Economic Review*, 104(1):149–82, 2014.

M. Angelucci, G. De Giorgi, and I. Rasul. "Resource Pooling Within Family Networks: Insurance and Investment". *mimeo, University College London*, 2015.

O. Attanasio and N. Pavoni. "Risk Sharing in Private Information Models with Asset Accumulation: Explaining the Excess Smoothness of Consumption". *Econometrica*, 79(4):1027–1068, 2011.

O. Attanasio and M. Szekely. "Wage Shocks and Consumption Variability in Mexico during the 1990s". *Journal of Development Economics*, 73(1):1–25, 2004.

J-M. Baland, I. Bonjean, C. Guirkinger, and R. Ziparo. "The Economic Consequences of Mutual Help in Extended Families". *mimeo, University of Namur*, 2015.

K. Beegle, R. Dehejia, and R. Gatti. "Child Labor and Agricultural Shocks". *Journal of Development Economics*, 81:80–96, 2006.

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

T. Besley. "Non-Market Institutions for Credit and Risk-Sharing in Low-Income Countries". *Journal of Economic Perspectives*, 9(3):115–127, 1995.

F. Bloch, G. Genicot, and D. Ray. "Informal Insurance in Social Networks". *Journal of Economic Theory*, 143(1):36–58, 2008.

T. Bold. "Implications of Endogenous Group Formation for Efficient Risk Sharing". *Economic Journal*, 119(536):562–591, 2009.

T. Bold and S. Dercon. "Insurance Companies of the Poor". *CEPR Discussion Papers*, 10278, 2014.

A. Chaudhuri, L. Gangadharan, and P. Maitra. ”An Experimental Analysis of Group Size and Risk Sharing”. *Unpublished*, 2010.

M. Comola and M. Fafchamps. ”The Missing Transfers: Estimating Mis-reporting in Dyadic Data”. *CEPR Discussion Paper No. DP10575*, 2015.

J. DeWeerdt, G. Genicot, and A. Mesnard. ”Asymmetry of Information within Family Networks”. *IZA Discussion Papers No. 8395*, 2014.

P. Dubois. ”Heterogeneity of Preferences, Limited Commitment and Coalitions: Empirical Evidence on the Limits to Risk Sharing in Rural Pakistan”. *Mimeo, Toulouse School of Economics*, 2006.

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

M. Fafchamps and F. Gubert. ”The Formation of Risk Sharing Networks”. *Journal of Development Economics*, 83(2):326–350, 2007.

M. Fafchamps and S. Lund. ”Risk-Sharing Networks in Rural Phillipines”. *Journal of Development Economics*, 71(2):261–287, 2003.

E. Fitzsimons, B. Malde, and M. Vera-Hernandez. ”Does increased interaction help consumption smoothing? Experimental Evidence from Malawi”. *mimeo, IFS.*, 2013.

E. Fitzsimons, B. Malde, A. Mesnard, and M. Vera-Hernandez. ”Nutrition, information and household behaviour: experimental evidence from Malawi”. *IFS Working Paper Series*, 2014.

A. Foster and M. Rosenzweig. ”Imperfect Commitment, Altruism and the Family: Evidence from Transfer Behavior in Low-Income Countries”. *The Review of Economics and Statistics*, 83(3): 398–407, 2001.

G. Genicot and D. Ray. ”Group Formation in Risk-Sharing Arrangements”. *Review of Economic Studies*, 70 (1):87–113, 2003.

M. O. Jackson, T. Rodriguez-Barraquer, and X. Tan. ”Social Capital and Social Quilts: Network Patterns of Favour Exchange”. *American E*, 102(5):1857–97, 2012.

P. Jakiela and O. Ozier. ”Does Africa Need a Rotten Kin Theorem? Experimental Evidence from Village Economies”. *Review of Economic Studies*, forthcoming.

C. Kinnan. ”Distinguishing Barriers to Insurance in Thai Villages”. *mimeo, Northwestern University*, 2014.

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

S. Lewycka, C. Mwansambo, M. Rosato, P. Kazembe, T. Phiri, A. Mganga, H. Chapota, F. Malamba, E. Kainja, M. Newell, G. Greco, A. Brannstrom, J. Skordis-Worrall, S. Vergnano, D. Osrin, and A. Costello. ”Effect of women’s groups and volunteer peer counselling on rates of mortality, morbidity and health behaviours in mothers and children in rural Malawi (MaiMwana): a factorial, cluster-randomised controlled trial”. *The Lancet*, 381:1721–35, 2013.

E. Ligon. ”Risk Sharing and Information in Village Economies”. *Review of Economic Studies*, 65(4): 847–864, 1998.

E. Ligon, J. Thomas, and T. Worrall. ”Informal Insurance Arrangements with Limited Commitment: Theory and Evidence from Village Economies”. *The Review of Economic Studies*, 69(1): 209–244, 2003.

B Mace. ”Full Insurance in the Presence of Aggregate Uncertainty”. *Journal of Political Economy*, 99(5):928–956, 1991.

M. Mtika and H. Doctor. ”Matriliny, Patriliny and Wealth Flow Variations in Rural Malawi”. *African Sociological Review*, 6(2):71–97, 2002.

A. Munthali. ”Adaptive Strategies and Coping Mechanisms of Families and Communities Affected by HIV/AIDS in Malawi”. *Draft paper prepared for the UNRISD project HIV/AIDS and Development*, 2002.

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

P. Peters, P. Walker, and D. Kambewa. ”Striving for Normality in a Time of AIDS in Malawi”. *The Journal of Modern African Studies*, 46(4):659–687, 2008.

K. Phiri. ”Some Changes in the Matrilineal Family System among the Chewa of Malawi since the Nineteenth Century”. *The Journal of African History*, 24(2):257–274, 1983.

A. I. Richards. *”African Systems of Kinship and Marriage”*, chapter ”Some Types of Family Structure Amongst the Central Bantu”, pages 207–251. Oxford University Press, 1950.

M. Rosenzweig. ”Risk, Private Information and the Family”. *The American Economic Review*, 78(2):245–250, 1988a.

M. Rosenzweig. ”Risk, Implicit Contracts and the Family in Rural Areas of Low-Income Countries”. *The Ec*, 98:1148–1170, 1988b.

M. Rosenzweig and O. Stark. ”Consumption Smoothing, Migration and Marriage: Evidence from Rural India”. *Journal of Political Economy*, 97(4):905–926, 1989.

O. Stark and R. Lucas. ”Migration, Remittances and the Family”. *Economic Development and Cultural Change*, 36(3):465–481, 1988.

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

R. Townsend. ”Consumption Insurance: An Evaluation of Risk-Bearing Systems in Low-Income Economies”. *Journal of Economic Perspectives*, 9(3):83–102, 1995.

J. Trinitapoli, S. Yeatman, and J. Fledderjohann. ”Sibling Support and the Educational Prospects of Young Adults in Malawi”. *Demographic Research*, 30:547–578, 2014.

X. Wang. ”Risk Sorting, Portfolio Choice, and Endogenous Formal Insurance”. *mimeo, Duke University.*, 2015.

F. Witoelar. ”Risk Sharing Within the Extended Family: Evidence from the Indonesia Family Life Survey”. *Economic Development and Cultural Change*, 62(1):65–94, 2013.

A Details of Model Simulation Calculations

In this section, we provide a step-by-step overview of the calculations that yield Figure 1 above. Given the specific parameter values associated with this particular example, groups of size, $N = 1, 2$ and 3 are found to be stable, while those of sizes, $N = 4 - 10$ are found to be unstable. The social planner randomly assigns households in a group of a specific size, N to stable subgroups of sizes $s_1, s_2, ...s_J$ in a manner so as to maximise total expected utility, while ensuring that all households are assigned to some stable sub-group. For groups of size, $N = 1, 2$ and 3, the social planner has no need to reassign households to stable sub-groups of a smaller size, s_j. Thus the average expected utility, and expected difference in marginal utility, for households in groups of these sizes can be recovered from Equations 4 and 5 by setting $\pi_s = 1$ for its group size and 0 for all other other stable group sizes and evaluating these equations at the optimal transfer. The calculated values are given in the Table 15 here.

Table 15: Expected Utility, and Expected Difference in Marginal Utility for Stable Groups, Example 1

Group Size	Expected Utility	Expected Difference in Marginal Utility
1	0.66206	0.15745
2	0.66377	0.03378
3	0.66857	0.03344

For groups of other sizes, we need to solve for the combination of stable sub-groups that maximises total expected utility when the social planner randomly assigns households to the stable subgroups. In this example, the optimal allocation of sub-groups for a group of size 4 is 1 sub-group of size 3 and 1 sub-group 1. Since households are randomly allocated into these sub-groups, each household has a $\frac{1}{4}$ chance of being in the sub-group of size 1 and $\frac{3}{4}$ of being in a group of size 3. The associated weighted average expected utility is thus

$$\frac{3}{4} * 0.66857 + \frac{1}{4} * 0.66206 = 0.66694$$

For a group of size 5, the optimal sub-groups are one of size 3 and one of size 2. Each household now has a $\frac{3}{5}$ chance of being in the group of size 3 and a $\frac{2}{5}$ chance of being in a group of size 2. The corresponding weighted average expected utility is

$$\frac{3}{5} * 0.66857 + \frac{2}{5} * 0.66206 = 0.66665$$

Note that the weighted average expected utility for a group of size 5 is lower than that for a group of size 4 because the probability of being in the higher utility sub-group of size 3 is higher in the latter case than in the former. This probability difference off-sets the increased expected utility from being in a sub-group of size 2 in the former case relative to being in one of size 1 in the latter case. Table 16 summarises these calculations for groups of sizes 4 - 10 in this example.

Table 16: Details of calculation for unstable groups, Example 1

Group Size	Prob. of being in stable subgroup of size:			Weighted Avg. EU	Weighted Avg. Expected Diff in MU
	1	2	3		
4	$\frac{1}{4}$	0	$\frac{3}{4}$	0.66694	0.06444
5	0	$\frac{2}{5}$	$\frac{3}{5}$	0.66665	0.03357
6	0	0	1	0.66857	0.03344
7	$\frac{1}{7}$	0	$\frac{6}{7}$	0.66764	0.05115
8	0	$\frac{1}{4}$	$\frac{3}{4}$	0.66737	0.03352
9	0	0	1	0.66857	0.03344
10	$\frac{1}{10}$	0	$\frac{9}{10}$	0.66792	0.04584

B Details of Simulations to Assess the Sensitivity of Parameter Estimates to Aggregate Extended Family Shocks

A concern is that our estimates might be biased since we are unable to suitably control for group-level shocks. We use simulations to assess the sensitivity of our parameter estimates to biases arising from this issue. Here we provide some details on the set-up of the simulations.

1. First, we generate a set of households and assign them to groups and villages. Villages contain multiple groups, and groups can span across multiple villages. Groups have different sizes, with the distribution of group sizes (total, and in the village) selected to match those found in the data.

2. We set the income process as follows: household income is composed of two components, a household-level component, $y_i = \{h_i, l_i\}$ and a group-level component, $y_g = \{h_g, l_g\}$. We select the values of h_i and l_i to be similar to those in our data. For the group-level shock, we set $h_g = 0$ and vary the values of l_g to be $\pi\bar{y}_i$, where $\bar{y}_i$ is the household's expected income. The probability of h_i is set to p, $0 < p < 1$; and that of h_g is set to π; $0 < \pi < 1$. Throughout, we set $p = 0.63$ and $\pi = 0.06$. The former probability is derived from our data, and the latter corresponds to the probability of a village-level aggregate shock in the data.

3. We extend the Genicot and Ray (2003) model to allow for common group level shocks (that are independent of the household-level shock), and given the values of h_i, l_i, h_g and l_g, and other parameters, compute the set of stable group sizes and derive the optimal transfer. We use the same consumption rule as in GR, and use the optimum transfers to calculate consumption under different states.

4. Given the set of stable group sizes, we allocate households in a potential group of size S to stable groups so as to maximise the total expected utility of the potential group. Since we assume the households are all homogenous, this amounts to a random allocation of households to stable groups.

5. We then randomly draw realisations of y_i for each household, and y_g for each group.

6. Given the stable group, and the realisations of y_i and y_g, we use the consumption rule computed in (3) above to assign consumption to each household.

7. We repeat (5) and (6) to attain a panel of shock and income realisations.

8. We then run specification 9, allowing first for the term μ_t^n to be a group-level dummy, and then for it to be a village-level dummy. We obtain the coefficients β_1, β_2 and β_3.

$$\begin{aligned}\Delta log(c_{ivt}) = \alpha_0 + \alpha_1 \Delta(crop_{ivt}) + \beta_1 \Delta crop_{ivt} * 1(NS_{g,iv} = 1) + \beta_2 \Delta crop_{ivt} * 1(NS_{g,iv} \geq 3) \\ + \mu_t^n + \Delta\epsilon_{int} \qquad (9)\end{aligned}$$

9. Repeat steps 4-8 100 times. Table 10 displays the results for different levels of the common group shock.