Managing Risk Heterogeneity in Risk-Sharing Groups: A Multi-Method Study on Risk Aversion and Solidarity

Eva Vriens and Vincent Buskens

Utrecht University

November 10, 2020

Abstract

A recent revival of organizations that provide insurance in risk-sharing groups (e.g., Broodfonds in the Netherlands, Friendsurance in Germany) brings us back to the age-old question of how risk-sharing groups can survive despite their risk heterogeneity. Data tells us that members within these groups vary substantially in their risk of needing support, essentially meaning that low-risk members pay more to support others than they receive support in return. How can heterogeneous risk-sharing groups realize stable participation rates? In a multi-method study, we examine the potential of risk aversion and solidarity as compensators for heterogeneity by comparing the results of a survey conducted among 5192 members of 230 Broodfonds groups to an online experiment with 430 British subjects of the Prolific platform. While we find that risk heterogeneity has a negative effect on participation in Broodfonds groups, our experimental manipulation of heterogeneity has no significant effect. Moreover, risk aversion does not predict participation in the field study nor in the initial decision to join in the experiment. In the first, it only predicts participation for low-risk members and in the latter, it only explains *continued* participation. Solidarity motives, finally, are a strong predictor in both settings. These results have important implications for theories on sharing risk (in which traditionally risk aversion is a crucial factor) and for understanding what may make risk-sharing groups successful in practice.

Keywords: risk-sharing, risk heterogeneity, solidarity, risk aversion, survey, experiment

1 Introduction

Over the past two decades, new mutual insurance organizations have been launched in many countries globally (see Vriens & De Moor, 2020 for an overview). In the Netherlands, for example, many self-employed workers have joined so-called Broodfonds groups (lit.: "bread funds") in which up to 50 members support each other in case long-term illness makes generating an income impossible. Other than regular insurance, which is paid in fixed premiums, the new organizations use risk-sharing arrangements: Contribution fees are used to build a common insurance fund of which the reserves not used to support members in need get redistributed at the end of the term. Thus, participants directly help each other, yet are all better off the lower the number of claims (redistribution then lowers the costs for all participants). This introduces an interdependency between members that could make participation particularly fragile the larger the heterogeneity in members' insurance risk (Platteau, 1997; Vogt & Weesie, 2004).

Risk heterogeneity implies an inequality where some group members have an above-average reliance on support from the risk-sharing group (Hegselmann & Flache, 1998). Others (the low-risk people) pay more to support these high-risk people than they receive support in return. Hence, participation is less attractive for low-risk people (the returns from redistribution are lower than they could be in a more homogeneous group), which is thought to hinder success (Breer & Novikov, 2015; Coate & Ravallion, 1993). While insurance companies solve this dilemma by decreasing premium rates for low-risk policyholders, many risk-sharing groups do not. Broodfonds groups, for instance, do not apply *ex ante* risk differentiation based on the premise that risk cannot be perfectly assessed and everyone should have access to support (Vriens & De Moor, 2020). From data collected among 230 Broodfonds groups (Vriens, Buskens, & De Moor, 2017), we know that there is substantial risk heterogeneity within these groups. What makes (low-risk) people willing to participate in heterogeneous risk-sharing groups? In this chapter, we study empirically to what extent risk aversion and solidarity can explain this pattern using a multi-method approach.

After all, Broodfonds groups are not exceptional. Although the new collective insurance organizations are covered with great curiosity and certain disbelief in European media (see, e.g., Curvers, 2019), sharing risk is the oldest means of creating security worldwide (Platteau, 1997). The earliest written accounts were found among the early modern guilds, where craftsmen came up with a pre-modern social security system for their members (see Epstein & Prak, 2008). Later, in the 19th century, mutuals were the most widespread—and arguably most successful—way of organizing insurance in most of Europe, the US, and Australia (Emery & Emery, 1999; Van

Leeuwen, 2016). While small-scale mutuals have slowly faded from the collective memory in these countries (as their tasks have largely been taken over by welfare states and large private insurance companies; Beito, 2000), they still form a key means to share risk among rural populations in sub-Saharan Africa (Lemay-Boucher, 2009), India (Ligon, Thomas, & Worrall, 2002), and Southeast Asia (Fafchamps & Lund, 2003).

This brings us to the gap between theory and practice, because while theoretical accounts often make the simplifying assumption of risk homogeneity (Coate & Ravallion, 1993; Ligon et al., 2002), in doing so they get rid of a crucial characteristic of real-world risk-sharing groups. Models that do relax the homogeneity assumption generally use risk aversion to explain why people may decide to share risk despite heterogeneity (Attanasio, Barr, Cardenas, Genicot, & Meghir, 2012; Breer & Novikov, 2015; Vogt & Weesie, 2004). Risk averse people prefer certain outcomes over uncertain gambles even if the expected value of the gamble is more profitable (Arrow, 1984). Applied to the context of sharing risk, it is hypothesized that this makes them willing to pay more to support others today to ascertain an income in the future (Platteau, 1997). While experiments corroborated that risk averse people were willing to share risk despite heterogeneity (Vogt & Weesie, 2006), many field studies show lower participation rates than would be expected from a risk aversion assumption (Platteau, De Bock, & Gelade, 2017). Hence, further empirical research is needed to shed light on the role of risk aversion—or more specifically, on whether risk-sharing groups are or are not considered an attractive, uncertainty-reducing alternative for risk averse people.

In small risk-sharing groups, participants do not only create financial security for themselves, but for others as well. This introduces a social aspect that could either strengthen resilience (when participants get motivated by solidarity from helping each other) or weaken it (when participants perceive the distribution of costs and benefits as unequal). Hence, the second aspect included in this study is solidarity, operationalized as pro-social motives directed towards members within the same risk-sharing group, i.e., beyond general motives of altruism (Baldassarri, 2015), from which personal utility (a good feeling) is derived as well (Gintis, Bowles, Boyd, & Fehr, 2005). We highlight this aspect particularly because it aligns with the vision of the new risk-sharing groups themselves. While their revival can mostly be attributed to a new and growing societal need—changing demographic structures, decreasing welfare states and increasing privatization of the insurance system have led to an increase in the number of people excluded from social insurances—the new organizations see a chance as well. Their mission statements describe a desire to go back to insurance as it was historically perceived: simpler, fairer, and kinder (Vriens & De Moor, 2020). In other words, solidarity has been given a central role, and empirical research is needed to find out whether that is justified.

In this chapter, we aim to gain insight into the dynamics of risk-sharing processes, particularly into mechanisms that drive people to participate or drop out, using two case studies. The first is a study among members of Broodfonds groups. Using survey data we study cross-sectionally how commitment (taken as predictor for future participation) relates to risk and insurance use (of both the individual and the group), risk aversion, and solidarity. While yielding insights in the role of risk aversion and solidarity for people embedded in real-world risk-sharing groups, the data only includes members and therefore lacks a counterfactual comparison. Therefore, we also conducted an online experiment among Prolific users. In a contextualized experiment we mimicked the structure of Broodfonds groups and studied under which circumstances people were willing to join and remain part of the group. This enabled us to truly test dynamics in participation processes, albeit tested in an abstract, artificial, and anonymized setting. Combined, the two case studies make up for each other's limitations (Buskens & Raub, 2008) and help to gain insight in how risks are shared, what the role of risk aversion and solidarity is, and more generally what characteristics allow for stable participation patterns in risk-sharing groups. If complementary results are found through both methods, we can—with more certainty—uncover the mechanisms behind the success of risk-sharing arrangements (Buskens, 2014).

2 Theory

The main theoretical framework to explain sharing risk and organizing mutual insurance is structured around the premise that the decision to participate represents a social dilemma (Fafchamps, 1992). While everyone could decide to save individually, this is more costly than saving collectively. However, saving collectively is surrounded by uncertainty. The main source of uncertainty—one absent in general public good dilemmas—is that people never know whether they, at some point, actually need support from the common fund. In fact, in the best case scenario they never do (Platteau, 1997). This makes participation in mutuals or risk-sharing groups particularly tricky, for the benefit (insurance in times of need) may never be obtained. Additional sources of uncertainty are not knowing how many others will decide to participate (common to all types of social dilemmas), how often they need support from the common fund (simultaneously), and whether they will remain members to reciprocate the favor in the future (Vriens, Buskens, & De Moor, 2019).

The classic risk-sharing model starts from a situation of symmetry. By assuming homogeneous risks (generally operationalized as equal opportunities of ending up with a high or a low payoff), they explain how long-term stable risk-sharing arrangements can emerge among rational, self-interested people (Coate & Ravallion, 1993; Fafchamps, 1992). Without such assumptions, adverse selection comes into play. Adverse selection describes the phenomenon that risk-sharing groups (or insurances in general) attract an above-average number of high-risk members (Akerlof, 1970). It implies that any rational actor, who decides whether or not to participate in a risk-sharing arrangement based on cost-benefit calculations, opts out of arrangements in which their risk is lower than the group average (Genicot & Ray, 2003; Ligon et al., 2002). When risk-sharing groups attract in particular high-risk people, participation becomes more expensive for everyone involved and the common fund may be insufficient to support everyone (i.e., relatively speaking, payouts from the common fund are needed more often).

This does not mean, however, that any risk heterogeneity inherently means the risk-sharing arrangement will fail. Hence, let us rephrase the theoretical argument from a comparison between homogeneous and heterogeneous risk-sharing groups (where heterogeneous risk-sharing groups are always worse off) to one that explores what extent of heterogeneity is acceptable. An early simulation model of Hegselmann and Flache (1998) studied how networks of mutual support can evolve in a world inhabited by rational egoists that need help with different probabilities (i.e., risk heterogeneity) and choose their risk-sharing partners endogenously in opportunistic ways. The model predicts that while people are, to some extent, willing to engage in unequal risk-sharing relations, they prefer partnerships that are as equal in risk as possible. In a different simulation study we predicted, for partnerships of $N > 2$ people and populations that vary in the degree of risk heterogeneity, whether utility-maximizing agents are willing to participate given specific risk aversion and solidarity traits (Vriens & Buskens, 2020). Our simulation results likewise showed that while stable participation patterns can emerge under various degrees of heterogeneity, in general group-level participation rates are lower the more heterogeneous the group. This effect was even stronger when agents did not know the actual risk distribution but based their estimate of the group-level risk on inferences from earlier support requests.

Experimentally, Vogt and Weesie (2006) found that risk heterogeneity decreases willingness to engage in dyadic risk-sharing relations, but that risk-sharing relations emerge nonetheless. Finally, Tausch, Potters, and Riedl (2014), found that when risk profiles are common knowledge, high-risk people were more likely to participate in 3-person risk-sharing groups than low-risk people and sharing risk was more likely to succeed in homogeneous rather than heterogeneous groups. Hence, driven by an interest to compare different degrees of heterogeneity, our baseline hypothesis is the following:

H1: *The lower the risk heterogeneity of a risk-sharing group, the larger the likelihood that members remain part of the risk-sharing group.*

The most popular explanation for why low-risk people participate in risk-sharing groups (or take out an insurance) is that people do not mind the extra costs if that guarantees them a safety net in the future. Called risk aversion (Arrow, 1971), the main idea is that not joining the risk-sharing arrangement introduces a gamble where with some probability you end up with nothing. The aversion towards this probability makes low-risk people willing to join such an arrangement even with high-risk people. Several theoretical variations on risk aversion exist, such as loss aversion (Kahneman & Tversky, 1979) or ambiguity aversion (Elabed & Carter, 2015), but we follow the majority of risk-sharing theories (Breer & Novikov, 2015; Coate & Ravallion, 1993; Delpierre, Verheyden, & Weynants, 2016; Laczó, 2014; Vogt & Weesie, 2004) and stick to the notion of plain risk aversion.

The positive effect of risk aversion to counter risk heterogeneity is corroborated experimentally (Vogt & Weesie, 2006). In field studies, risk aversion is rarely measured explicitly, but there is some evidence that participating in risk-sharing arrangements is lower than would be expected under risk aversion (Platteau et al., 2017), suggesting that risk-sharing arrangements may not always be an uncertainty-reducing alternative. To assess whether the theoretical role assigned to risk aversion holds in practice, we therefore test the following individual-level hypothesis:

H2: *The higher members' risk aversion, the larger the likelihood that they remain part of the risk-sharing group.*

The second individual-level explanation we put to a test is the role of solidarity motives. This requires a bit more clarification, as solidarity is used to describe different mechanisms in different studies. In the seminal work of Fafchamps (1992), for instance, solidarity networks are used to describe the age-old risk-sharing networks in place in preindustrial societies. Fafchamps refers to solidarity as a form of mutual insurance, where the person receiving assistence is expected to help others in return—without specifying how much help is warranted. In other words, solidarity is used to describe self-regarding direct or indirect reciprocity. In this case, however, we refer to solidarity as a type of pro-social motive that extends cooperation beyond basic reciprocal and reputation-based cooperation (Chaudhuri, 2011; Ostrom, 1990). It involves (to some extent) utility derived from helping others without expecting help in return (Fehr & Schmidt, 1999), and as such may invoke willingness to participate in risk-sharing groups when self-interested people would not.

While often discussed as a general personality trait (e.g., altruism, Levine, 1998), we believe that in the context of risk-sharing groups the distinction by Baldassarri (2015) is more applicable. In the experiment of Baldassari, solidarity (pro-social motives towards members of one's group) is found to lead to higher cooperation rates than altruism (pro-social motives in general). Given that risk-sharing groups do not necessarily need people to be pro-social in general, but rather willing to pay the costs to support others within the group (Vriens & Buskens, 2020), this particular notion of pro-social motives seems most applicable, and also best in line with the ideology of the new mutuals themselves. Experimentally, solidarity—or rather social preferences in general—are found to increase participation in informal risk-sharing arrangements when contrasted to the introduction of formal insurances (Lin, Liu, & Meng, 2014; Lin, Meng, & Weng, 2019). Finally, studies about historical mutuals of the 19$^{\text{th}}$ century often place solidarity motives central to explanations of their success (Harris, 2012).[1] Hence, our third hypothesis reads:

> **H3:** *The higher members' solidarity, the larger the likelihood that they remain part of the risk-sharing group.*

The above hypotheses all concern generalized predictions. There are, however, good reasons to believe that the effects of risk aversion and solidarity may depend on the group context—and thus on risk heterogeneity. That is, when risk averse people are looking for the solution that provides them more security, it may be questioned whether all risk-sharing groups are equally capable of providing this (Platteau et al., 2017). After all, risk-sharing groups introduce many uncertainties as well (e.g., how many people need support and how often, and whether the fund will be sufficient to help everyone). These uncertainties increase with heterogeneity. Likewise, if we hypothesize on the role of solidarity as pro-social behavior towards the members of the risk-sharing group (rather than pro-social behavior in general), it may matter who these people are (Vriens et al., 2019).

Through simulations, we showed that while risk aversion and solidarity can advance participation despite risk heterogeneity, they are more effective in doing so when heterogeneity is lower (Vriens & Buskens, 2020). While this may seem contradictory at first—after all, these two factors generally enable participation under inequality—this result should be interpreted with an eye on how these factors were operationalized to increase the motivation to participate. Taking solidarity as example, any member with at least some degree of solidarity motives is

[1] It should be noted that while case studies of historical mutuals often mention solidarity (and social norms, social control) as an important reason for success, this is discussed more in relation to reducing moral hazard than for countering adverse selection. Moral hazard (when insured people increase risky behave or exaggerate insurance needs, Arrow, 1971) falls beyond the scope of the current investigation.

willing to pay higher costs to help other group members (solidarity compensates for these costs). However, it will only compensate up to some degree. If, in more heterogeneous groups, the cost fluctuations of supporting others are larger and higher over time, the probability increases that the degree of solidarity (or risk aversion) is insufficient. Hence, to assess the dynamics of risk aversion and solidarity we derive the following hypothesis:

> **H4:** *The higher the risk heterogeneity, the weaker the positive effect of* ***(a)*** *risk aversion and* ***(b)*** *solidarity on the likelihood that members remain part of the risk-sharing group.*

Finally, to further investigate the dynamics and interdependencies between the decision-making of different group members, we study how the behavior of other group members affects individual decision-making. Collective action theories state that people are willing to participate only as long as a sufficient number of others also participate (i.e., when the number of participants is above a certain threshold, Granovetter, 1978). This goes for the willingness to start to cooperate in the first place, but remains just as true in deciding whether or not to continue to cooperate. Any collectively established cooperation arrangement is fragile to small (or rare) disturbances, which can cause a downward cascade of cooperation (Ostrom, 2005).

In the context of risk-sharing groups, fluctuations in the number of group members requesting support could be such disturbances. When suddenly several group members need support, the costs of participation go up. This could mean that for one or a few members, participation may no longer be interesting. If they then choose to opt out of the risk-sharing arrangement, this drives the remaining members to reevaluate whether the risk-sharing arrangement is worthwhile given a smaller number of people with whom risk is shared. Especially assuming that low-risk members are more likely to opt out after sudden or temporary cost increases, a decrease in the total number of members also means an increase in the average participation costs per round. Our earlier simulation study also showed that dropout by one (or several) members could easily set in motion withdrawal cascades (Vriens & Buskens, 2020). Our final hypothesis thus reads:[2]

> **H5:** *Sudden increases in support requests and dropouts by one (or several) member(s) set in motion cascades such that other members are less likely to remain part of the risk-sharing group either.*

[2] This hypothesis cannot be tested with the cross-sectional data of the Broodfonds groups (Study 1) and will therefore be tested with the experimental data only (Study 2).

3 Study 1: Broodfonds Survey

3.1 About Broodfonds

Broodfonds is a collective in which self-employed workers in the Netherlands (who are excluded from welfare benefits regarding retirement and disability) organize their own financial safety net. Everyone pays a monthly contribution that is used to create a common fund from which members suffering from long-term illness can be supported. At the time of writing (October 2020), there are 595 Broodfonds groups in the Netherlands with an average of 45 members per group (26,900 in total). Each group, while based on the same organizational framework, is registered as an independent association (i.e., not formally recognized as an insurance). Governed by board members that are sequentially chosen and appointed from their own member base, they share the risk of income loss due to illness on the basis of trust. Sick members do not need to prove illness through official doctor's notes or house visits, but are asked to update the group on their progress towards recovery.

The first Broodfonds group started in 2006. Subsequently, three members of this group started a new cooperative called BroodfondsMakers that, as an umbrella organization, supports other self-employed workers to start their own Broodfonds group. These new groups have to adhere to the basic organizational framework to carry the name Broodfonds. The number of Broodfonds groups grew rapidly afterwards, from 18 by the end of 2012 to 230 by February 2017 (reference date for the data collection) to 595 by October 2020.

Common membership restrictions include requirements such as the necessity to work or live in the same municipality and to be introduced by one or two existing group members. Risk is never a restriction: as a policy, Broodfonds groups do not differentiate based on (predicted) risk. In their opinion risk cannot be estimated precisely and everyone deserves basic support levels. Yet since groups are formed endogenously, it could be that risk does factor in implicitly when people decide with whom to join a group (e.g., there are a few groups that consist entirely of people working in the same sector). As such, substantial variation may exist between the different groups in terms of risk heterogeneity.

3.2 Data

We rely on an online survey conducted among all 10,331 members of the 230 Broodfonds groups that were established before February 2017 (Vriens et al., 2017). This survey, administered between May 10 and June 14, 2017, asked about personal characteristics, membership motivations, and use of the mutual fund. The chairpersons of the 230 groups were asked to fill in a second survey with questions about organizational properties and support. 5,192 respondents filled in the member questionnaire (50.7%). The organization questionnaire was filled in for 196 of 230 groups (85.2%). The data collection has been approved by the Ethics Committee of the Faculty of Social and Behavioral Sciences of Utrecht University (reference number 17-042).

3.3 Methods

An obvious drawback of using a cross-sectional survey among members of existing risk-sharing groups to test theories about participation is, of course, that all respondents participate, meaning that strictly speaking there is no variation on the outcome variable. We can, however, look at people's intentions to continue membership in the future. One often-used factor to measure such intentions is commitment (Hauert, 2002; Kollock, 1994; Orbell, Schwartz-Shea, & Simmons, 1984; Vriens et al., 2019). Low levels of commitment signal a potential threat, as members are more likely to leave in the future; high levels make it more likely that members will continue to participate regardless of potential (internal or external) changes.

Commitment was measured through seven items (obtained from Meyer, Allen, & Smith, 1993) that covered both affective aspects (emotional attachment) and normative aspects (perceived obligations).[3] Example items are "*I tell others proudly that I am part of this Broodfonds*" (affective, Van der Lippe et al., 2016) and "*Even if it were to my advantage, I do not feel that it would be right to leave Broodfonds right now*" (normative, Jak & Evers, 2010). Responses to all questions were measured on a 7-point scale ranging from (0) "completely disagree" to (6) "completely agree", where higher scores reflect more commitment. Because the items form an adapted selection of the original scales we used Exploratory Factor Analyses to test their validity and construct a scale based on the factor loadings (which were above the 0.32 threshold for all items). The scale has a Cronbach's alpha reliability of $\alpha = 0.82$.

To predict commitment, we take proxies for risk aversion, solidarity, risk, benefit size, and time on the individual level, as well as risk heterogeneity, the number of support requests, and group size on the level of the mutual group.

Risk aversion was measured using the "staircase" method: a validated method of inferring risk preferences through iterative multiple price lists (Andersen, Harrison, Lau, & Rutström, 2006; Falk, Becker, Dohmen, Huffman, & Sunde, 2016). It presents respondents with five choices

[3] In organization research, continuance commitment is generally called upon as a third dimension of commitment. This dimension depends on external factors (i.e., the presence or absence of attractive alternatives), and does not measure individual efforts of making the organization successful. Because small mutuals generally arise when no (or few) alternatives are available, affective and normative commitment are most useful as proxies for success.

between a lottery and a sure payoff, where the lottery is constant (always a 50/50 gamble of 300 versus 0 points) and the size of the sure payoff varies conditional on the previous decision (i.e., a higher sure payoff if the subject chose the lottery and vice versa). For the survey, a 1:1 translation was used so that 300 points were presented as a hypothetical €300. The safe counteroffer ranged between €10 and €310 depending on the position on the staircase. After five questions, respondents ended up on position 0 to 31 on the staircase, where 31 is the most risk averse position and 0 the most risk seeking.

Respondents were also asked about their motivation for joining a Broodfonds group (with options ranging from financial, to organizational and social motives). We used three motives that capture social aspects ("*To be able to do something for other people*", "*Solidarity towards each other*", and "*Being part of a group*") to create a factor scale of solidarity. All items had factor loadings above the 0.32 threshold and the scale has a Cronbach's alpha reliability of $\alpha = 0.75$.

As proxies for individual risk we used two indicators: A self-rated health status (perceived risk) and whether or not the respondent received a benefit from the group in the twelve months preceding the survey. Perceived risk was measured by asking respondents to rate their current (mental and physical) health on a 7-point scale ranging from (0) "very bad" to (6) "very good".

Broodfonds groups use fixed combinations of contributions and corresponding threshold sizes, so we also used information about the size of the benefit in case they would need support. Options ranged from €500 to €3000 in steps from €250. The resulting variable was scaled by dividing the answer by 100.

Time, finally, was included as the number of years that the Broodfonds group exists before 2017 (the year the survey was conducted). In addition, we included a 'deviation' variable indicating for every member of the group whether they joined Broodfonds earlier or later than this group's starting date. This deviation is positive if a group member joined an existing group and negative in case a group member switched to a different group. The latter occurred, for instance, when a group got bigger and decided to split into two separate groups.

For the risk of the Broodfonds group we again used two indicators. We calculated, for each Broodfonds group, the standard deviation of the self-rated perceived risk question as an indicator of risk heterogeneity. Moreover, we used the sum of the number of respondents who indicated that they received a benefit divided by the total number of members as an indicator of realized risk. Finally, we included the number of members of each Broodfonds group.

To compare our analyses to earlier analyses on the same data (Vriens et al., 2019) and to prevent that effects can be attributed to other relevant factors used before, we added several additional variables as controls. These include age, gender, trust in other group members, the

Table 1. Descriptive statistics for member- ($N = 3570$) and group-level ($N = 230$)

Variable	**M**	**SD**	**Min**	**Max**
Members				
Commitment	3.90	0.93	0.00	5.43
Risk aversion	19.70	6.98	0.00	31.00
Solidarity	3.44	0.88	0.00	4.70
Perceived risk	0.94	1.02	0.00	6.00
Received benefit	1.05	0.23	1.00	2.00
Benefit size	17.13	5.74	5.00	30.00
# Years Broodfonds − # Years member	0.65	1.10	-5.00	10.00
Gender: Female	1.47	0.50	1.00	2.00
Age	49.20	8.70	21.00	74.00
Trust	0.03	0.97	-4.54	1.38
Total degree	7.07	9.53	0.00	100.00
Strong tie degree	41.95	34.69	0.00	100.00
Groups				
Risk heterogeneity	0.98	0.28	0.44	1.83
Relative number of benefits	0.05	0.05	0.00	0.25
Total number of members	44.83	6.75	21.00	53.00
# Years Broodfonds group	2.59	1.43	0.00	11.00
Dense network	0.14	0.16	0.00	1.00
Clustered network	0.46	0.19	0.00	0.95
Star network	0.24	0.13	0.00	0.67
Sparse network	0.16	0.17	0.00	0.80

number of (strong) ties to other group members, and the network structure of the different Broodfonds groups (dense, clustered, star-shaped or sparse). The latter were measured as cognitive social network structures (Krackhardt, 1987) and depict the proportion of members of each group that considered that network structure to be the best description for the relations within their group.

Table 1 shows the descriptive statistics of all variables. There are slightly more men than women in the sample (54%) and members are on average 49 years old. Most respondents feel relatively healthy and the average realized risk (in terms of who received a benefit in the last twelve months) in the group was 5%.

We used multilevel OLS regression analyses to predict commitment using a two-step approach: First a model (M1) with all direct effects, followed by a model (M2) that included interactions of risk aversion and solidarity with risk heterogeneity. The models were compared with reference to the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), which consider one model superior over the other once the difference exceeds the threshold of 10 (Burnham & Anderson, 2003).

3.4 Results

The starting point for this study was the risk heterogeneity in Broodfonds groups. Using self-assessed risk perceptions aggregated to the group-level, the average risk per Broodfonds group was 0.94 (almost 1 point above the 7-point scale minimum of 0), indicating that on average Broodfonds members of all groups consider themselves to be healthy. However, the standard deviation ranges from $(0.44 - 1.83)$, with an average of 0.98 (SD = 0.28). This indicates that there is substantial variation both between and within groups in the individual risk perceptions. The correlation with member commitment rates is $\rho(228) = -.214$ ($p = .001$) on the group level and $\rho(4939) = -.053$ ($p < .001$) on the member level. That is, in groups were heterogeneity is higher, commitment of members seems to be lower.

Table 2 reports the results of the multilevel regression analyses, to test which relations hold significantly while controlling for group-level clustering and the relevant individual and group-level covariates. In general, most variance occurs on the member level. Only 2.7% of the variance in the empty Model M0 is shared on the group level. Compared to the empty model, Model M1 explains 47% of the variance on the member level and 67% of the variance on the group level.

As already expected from the bivariate correlation, Model M1 shows a negative effect of risk heterogeneity on commitment. In groups where the standard deviation of average perceived risk is 1, commitment is 0.147 points lower. This supports Hypothesis 1. The effect of risk aversion is not significant, which implies no support for Hypothesis 2. Solidarity, on the other hand, has a strong positive effect on commitment, with 1 additional point on the solidarity scale equating 0.427 points extra on the commitment scale. This provides support for Hypothesis 3. Finally, Model M2 adds the interaction effects between risk heterogeneity and risk aversion and solidarity. It does not explain additional variance, nor does it improve the model fit (the AIC and BIC scores are worse). Neither one of the interaction effects is significant, which means that the effects of risk aversion and solidarity do not depend on group-level heterogeneity. Hypothesis 4 is not supported.

Hypothesis 5 (about cascades in participation decisions) cannot be tested with the survey data, because all respondents in the cross-sectional sample are participating. However, commitment does not seem to be affected by factors that could be indicators for member interdependencies: Neither the relative number of benefits nor group size are significantly related to commitment. Interestingly, individual perceived risk does not affect commitment either. One would expect people with higher risk to be more committed (since they are more likely to need support from the risk-sharing group in the future), but this does not follow from the data. It

Table 2. Multilevel OLS regression analyses on commitment ($N = 3570$)

	(M0)		(M1)		(M2)	
Members						
Intercept	3.894^{***}	(0.017)	1.662^{***}	(0.143)	3.049^{***}	(0.131)
Risk aversion			0.003	(0.002)	0.003	(0.002)
Risk aversion × Risk heterogeneity					0.001	(0.006)
Solidarity			0.427^{***}	(0.014)	0.427^{***}	(0.014)
Solidarity × Risk heterogeneity					0.080	(0.049)
Perceived risk			−0.002	(0.012)	−0.001	(0.012)
Received benefit			0.175^{**}	(0.056)	0.169^{**}	(0.056)
Benefit size			0.003	(0.002)	0.003	(0.002)
# Years Broodfonds − # years member			−0.021	(0.012)	−0.021	(0.012)
Gender: Female			0.007	(0.024)	0.008	(0.024)
Age			0.015^{***}	(0.001)	0.015^{***}	(0.001)
Board member			0.135^{**}	(0.042)	0.135^{**}	(0.042)
Trust			0.327^{***}	(0.013)	0.327^{***}	(0.013)
Strong tie degree			0.005^{***}	(0.001)	0.005^{***}	(0.001)
Total degree			0.002^{***}	(0.000)	0.002^{***}	(0.000)
Groups						
Risk heterogeneity			-0.147^{**}	(0.054)	-0.148^{**}	(0.054)
Relative number of benefits			0.303	(0.303)	0.304	(0.303)
Total number of members			−0.001	(0.002)	−0.001	(0.002)
# Years Broodfonds			-0.034^{**}	(0.012)	-0.034^{**}	(0.012)
Dense network			0.134^{**}	(0.045)	0.135^{**}	(0.045)
Clustered network			0.026	(0.034)	0.026	(0.034)
Star network			0.034	(0.037)	0.034	(0.037)
Random intercept group	0.024		0.008		0.008	
Residual variance	0.843		0.444		0.444	
Intraclass correlation	0.027		0.019		0.018	
Log Likelihood	−6,645.275		−3,707.143		−3,712.055	
Akaike Inf. Crit.	13,296.550		7,458.285		7,472.110	
Bayesian Inf. Crit.	13,316.070		7,594.252		7,620.437	

Note: $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$

could be that this effect is captured partly by the variable indicating whether or not the member received a benefit. We do find members that received a benefit in the last twelve months to be more committed than those who did not. Since there is a moderately strong correlation between perceived risk and whether a benefit was obtained ($\rho_s = 0.266, p < .001$), it could be that perceived risk is captured in this variable as well. Those who received a benefit might perceive their risk to be higher, thinking they might need support again in the future. At the same time, it could be a sign against moral hazard or free riding: there is no incentive for people to drop out once they obtained their benefit. Instead, they might feel like it is up to them to reciprocate the favor in the future.

Finally, commitment is lower in older groups than in younger groups. There is no way to disentangle whether this is a cohort effect (i.e., older groups where less committed to begin with) or an age effect (i.e., commitment decreases over time), which makes it important to keep monitoring commitment levels of members over time. The remaining variables affect commitment in the expected directions. Like in Vriens et al. (2019), women, elderly people, people with higher trust levels, and people with more (strong) ties to other group members are more committed. Moreover, commitment is higher in dense networks than in networks with few to no relations among members.

Since high-risk members have immediate individual incentives to be members of the risk-sharing group, for them risk-aversion or solidarity should hardly play a role. Therefore, these effects might be perceived more clearly if we repeat the analyses for low-risk members only (the ones most likely to withdraw). The low-risk subgroup was defined by selecting within each Broodfonds group those respondents whose individual perceived risk was lower than or equal to the group's average (resulting in a total of $N = 1973$ respondents). In general, most results remain stable. There are, however, two important differences. The negative effect of risk heterogeneity is stronger for the low-risk subgroup ($b = -0.224$, $p = 0.002$) compared to the full sample ($b = -0.147$, $p = 0.007$), indicating, as the theory would predict, that the effect of heterogeneity is mainly detrimental for low-risk members. Secondly, for the low-risk subgroup we do find a positive effect of risk aversion ($b = 0.004$, $p = 0.038$), which indicates that the mechanism underlying risk aversion (reducing uncertainty) is relevant only when perceived risk is low(er). High-risk members may be more certain to at some point need support in the future (or perceive their risk to be high because they are already receiving support), so they are committed regardless of their general risk aversion. Detailed results of the subgroup regression analyses are displayed in **??** in **??**.

4 Study 2: Online Experiment

4.1 Experimental Design

Because the survey data of the members of Broodfonds groups only yielded intentions to continue membership and cross-sectional correlations, we also designed and conducted an online experiment. The experiment was contextualized and largely followed the set-up of Broodfonds groups. Subjects were asked to imagine working as freelancers for 20 rounds (which were explained to reflect months). Every month they worked on an assignment that earned them a steady income of 900 points (£6). When they got sick, they could not finish the assignment and would have no income. In the first round, they were invited to share the risk of getting sick and earning nothing with up to nine others. If they accepted, they would pay a contribution of 300 points

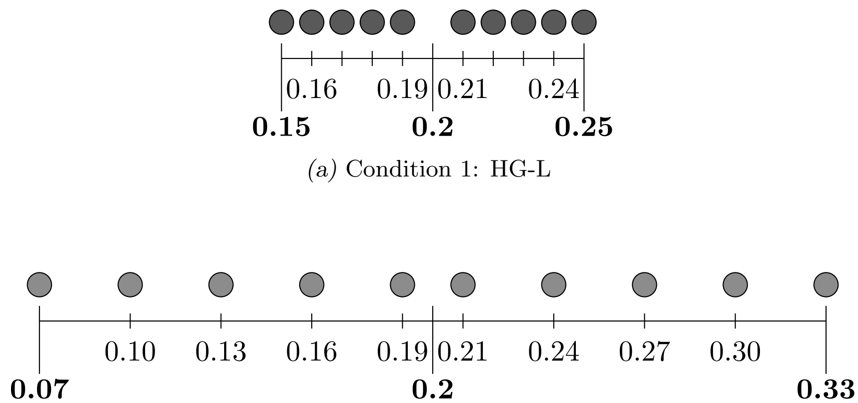
(£2) that would be used to pay sick members of the risk-sharing group. A sick member would receive a benefit of 750 points (£5)—unless the sum of sick members × 750 points exceeded the sum of contributions. In that case the sum of contributions would be equally divided over the sick members.

At the end of every round, the share of contributions that was not needed to pay benefits got redistributed across all members. As such, subjects that did not join the risk-sharing group earned 900 or 0 points depending on their health status. Healthy members of the risk-sharing group earned between 900 points (when no one got sick and the entire contribution fee could be returned) or 600 points (when the entire contribution fee was needed to pay sick members). Sick members, lastly, earned between 675 points (750 - 300 + 225) if they were the only sick member in a group of 10 and 0 points (-300 + 300) if all members would be sick simultaneously (the latter being very unlikely).

At the end of the experiment, one of the 20 rounds would be randomly drawn and added to the subject's earnings. Therefore, we opted for a higher contribution fee of which the remainder would be returned every round, instead of saving contribution fees in a common fund over time (as is done by Broodfonds groups). We chose this approach following earlier risk-sharing experiments (Charness & Genicot, 2009; Lin et al., 2014). Paying subjects the sum of all earnings in all rounds would imply that they saved their entire income and subjects would be incentivized to care about the final sum. In practice, however, most of one's income goes to paying rent and utilities, which is why illness (and the resulting lack of income) poses a serious threat. By randomly drawing one period for payout, subjects were incentivized to smooth consumption over periods.

Subjects were told that other members in their risk-sharing group might not have the same risk of getting sick. Yet while they knew their own risk, they did not know the risk of other members nor whether it was higher or lower than theirs. This they had to infer from the realizations of illnesses over time. Using a between-subjects design, we systematically compared two treatment conditions in which the subject population has the same average risk ($\bar{p} = 0.2$), but varies in the internal degree of risk heterogeneity. Participants play either the low (HG-L) or the high (HG-H) heterogeneity condition, with risks ranging between $0.15 - 0.25$ and $0.07 - 0.33$, respectively. The probability distributions in these conditions are displayed in Figure 1, where each circle represents the risk probability for one experimental subject.

For each low risk heterogeneity treatment, we ran a high heterogeneity treatment twin in another session that uses the same realization of sickness events over the treatment's round, albeit divided over different players. That is, while the risks are assigned randomly and it is



(a) Condition 1: HG-L

(b) Condition 2: HG-H

Figure 1. Probability distributions for the two risk heterogeneity treatments

also random which player gets sick in each round, we made sure that in all groups the same number of total sickness events occurred (i.e., $N\bar{p}T = 10 \times 0.2 \times 20 = 40$ events). These 40 events were randomly divided over 20 rounds and for each division we ran an HG-L and an HG-H session, in which the sickness events were randomly assigned to the participants based on their risk weight. This way of systematically controlling the total number of events reduces the noise that is introduced by translating probabilities to realizations and provides a cleaner test of the differences between the heterogeneity conditions given our small sample size.

In the first round, participants were asked whether they wanted to join the risk-sharing group. If they joined, in all subsequent rounds they would be asked whether they wanted to *continue* to be part of the risk-sharing group. If they would not join (round 1) or would decide to opt out (all later rounds) they did not get a chance to return at a later stage. Subjects who do not participate only saw their own realization of every round (i.e., whether they got sick or not), but no longer had a choice to make.

Subjects who did join were informed about the realizations in their risk-sharing group every round: whether or not they themselves got sick, how many members the risk-sharing group has, what the health status of the other members is, how much money was needed to pay sick group members, how much of the contribution was returned, and how much they would earn if that round were drawn for payment. Before the start of the experiment, subjects had to read the instructions (which were the same for all subjects regardless of the treatment), were presented with example screenshots of what the decision situation could look like in different rounds, and had to answer a series of questions to verify that they had read and understood the instructions correctly. The instructions, example screenshots, and questions are included in **??**.

4.1.1 Risk Preferences and Solidarity

Before the risk-sharing game we present the subjects with several other decision scenarios to measure risk preferences and solidarity. For risk preferences we use the same "staircase" method (Andersen et al., 2006; Falk et al., 2016) that we used in the survey among Broodfonds members. In five decision situations, they had to choose whether they would prefer a lottery (always a 50/50 gamble of 300 versus 300 points) or a fixed amount (ranging from 10 to 310 points).

As an indicator for solidarity we used the Social Value Orientation Slider Measure (SVOSM; Murphy, Ackermann, & Handgraaf, 2011), which is found to outperform its competitors (the Triple-Dominance Measure and the Ring Measure) in most respects (Bakker, 2019). The SVOSM presents the subject with 15 dictator games that vary in the conversion rates of points allocated to the decision maker (between 50 and 100 points) and the recipient (between 15 and 100 points). The subjects knew that the recipient is a real person. Choices are made in real time and they have to wait for the other person to make a decision before they continue. They do not know, nor were they informed afterwards, who this other person is. The average score of all points assigned to oneself can be translated into an overall score, the SVO angle, that indicates the degree of prosociality.

We randomly choose the outcome of one lottery and dictator game to add to the payoff of participants, to avoid wealth effects where choices to the current decision are influenced by the outcome of the previous decision (Azrieli, Chambers, & Healy, 2018; Harrison & Elisabet Rutström, 2008).

4.2 Data and Procedure

The experiment has been approved by the Ethics Committee of the Faculty of Social and Behavioral Sciences of Utrecht University (reference number 20-229). The experiment was programmed using o-Tree (Chen, Schonger, & Wickens, 2016) and conducted with UK-based participants recruited from Prolific. Participants signed up for a session a day in advance. When they started the session, they first had to read and agree to a consent form, which informed them about the confidentiality and anonymity of their data, the data storage term (10 years), the ethical approval, and the open access of the anonymized data. They were instructed that the study consisted of four parts, of which parts two and three were played interactively (with 1 and 9 other persons, respectively). In the instructions we stressed that these were real people and not simulated bots and that they should keep their attention to the screen to not keep other people waiting.

The study consisted of the staircase method (Part 1), the SVOSM (Part 2), the risk-sharing game (Part 3) and a survey (Part 4). As soon as people had given consent they could start with Part 1. Parts 2 and 3 started when enough people arrived at these stages. If not enough people showed up to form a group, after 15 minutes the waiting participants were redirected back to Prolific and received a show-up fee of £5.

For Part 1, we used an oTree implementation of the staircase method that was programmed by Holzmeister (2017). The implementation presents five sequential lottery questions, where the subject is asked to choose between option A (the lottery) or option B (the sure payoff) and keeps track of the progress by means of a progress bar. For Part 2, we programmed a continuous version of the SVOSM in oTree, where the subjects can choose a division of points between themselves and the receiver using a slider. For Part 3, we programmed the formation of groups *after* the subjects had finished reading the instructions and answering questions, so that people would not be stuck waiting for others. Part 4, finally, consisted of a survey. For all subjects that participated in the risk-sharing group for at least one round, the survey started some questions about their motivation for joining the risk-sharing group. All subjects were asked about several demographics.

At the end of Part 4, subjects were informed about their earnings. For each of Parts 1, 2, and 3, one round was randomly chosen and added to the subject's earnings. The exchange rate for all points earned during the experiment was 150 points = £1. Combined with a show-up fee of £2, this meant that subjects could earn between £5 and £11.40 for their participation in the experiment.

The data was collected in twelve sessions conducted between June 5 and June 11, 2020. In total, 525 people showed up for these sessions (68% of those who signed up), of which 430 people could be assigned to a group for Part 3 of the experiment. Of those 430 people, 424 completed the entire study. They needed between 35 and 95 minutes to complete the experiment and earned between £5 and £10.63 (M = £9.04, SD = £1.69).[4]

4.3 Methods

As dependent variable, we store for all participants in all groups and all rounds whether or not they participated in the risk-sharing group. As main predictor variables, we take the treatment condition (HG-L versus HG-H) and the subject's score on risk aversion and solidarity (the SVO angle). Risk aversion is measured the same as in Study 1. The SVO angle is computed as

[4]For 31 subjects, the final earnings ended up below £5, but their earnings were raised to the base fee of £5.

Table 3. Descriptive statistics per treatment

Treatment	**Variables**	**M**	**SD**	**Min**	**Max**	**N**
Subjects						
HG-L	Risk	0.20	0.03	0.15	0.25	207
HG-H		0.20	0.08	0.07	0.33	216
HG-L	Risk aversion	20.05	6.07	0.00	31.00	207
HG-H		19.66	6.65	0.00	31.00	216
HG-L	Solidarity	32.73	9.09	11.62	46.60	207
HG-H		34.07	8.04	11.62	49.81	216
HG-L	Gender: Female	0.75		0	1	207
HG-H		0.75		0	1	216
HG-L	Gender: Other (not disclosed)	0.01		0	1	207
HG-H		0.00		0	1	216
HG-L	Age	3.83	1.47	1.80	7.60	207
HG-H		4.08	1.49	1.80	7.60	216
HG-L	Game theory	0.09		0	1	207
HG-H		0.06		0	1	216
Groups $\times$ Rounds						
HG-L	Total members $t-1$	6.60	2.02	2	10	420
HG-H		6.56	2.20	0	10	440
HG-L	New withdrawals $t-1$	0.93	0.44	0	3	420
HG-H		0.95	0.46	0	4	440
HG-L	# sick members RSG $t-1$	1.26	1.23	0	6	420
HG-H		1.26	1.28	0	6	440
Groups $\times$ Subjects $\times$ Rounds						
HG-L	Participate	0.64		0	1	4140
HG-H		0.64		0	1	4320
HG-L	Sick $t-1$	0.19		0	1	4140
HG-H		0.19		0	1	4320

SVO° $= \arctan(\frac{\overline{A}_o - 50}{\overline{A}_s - 50})$, where $\overline{A}_o$ is the average amount the subject assigned to the recipient and $\overline{A}_s$ is the average amount kept. A prosocial subject with inequality aversion would yield an angle of 37.38°; a perfectly consistent individualist yields an angle between −7.82° and 7.82° (Murphy et al., 2011).

Other predictors included are the risk probability assigned to each subject and, about the round before, whether they were sick, the total number of members, the number of members who dropped out, and the number of sick members in the risk-sharing group. As control variables, we include the subject's gender and age (divided by 10) and whether the subject indicated to have knowledge about game theory.

Table 3 shows the descriptive statistics of all variables. The majority of participants in the experiment was female (75%), the average age of subjects was 39. Only 32 respondents indicated to have any knowledge of game theory. The average subject was considerably risk averse (M = 19.88) and prosocial (M = 33.34).

To analyze the data, we started by comparing—on the level of the groups—how many subjects participated per round. Each HG-L group was matched to its HG-H twin (based on the total number of sickness events per round). We have data for 21 HG-L groups and 22 HG-H groups. Because the 22$^{\text{nd}}$ HG-H group does not have a twin, we excluded this group from the group-level analysis. Note that while the other 21 groups are paired, within the risk-sharing groups there might still be differences in the number of sick members, because the subject that got sick in a round may not be part of the risk-sharing arrangement. Still, the starting characteristics of these groups are comparable, so we used a paired samples t-test to compare each group $\times$ round combination of the two treatments.

Subsequently, we used Event History Analyses (EHA) on the subject $\times$ round level to predict withdrawal (the event) from risk heterogeneity, risk aversion, solidarity, and all other model parameters. EHA is commonly used to analyse time-to-event data. The focus is on the modelling of event transition (i.e., from participating to not participating) and the time it takes for the event to occur. The benefits of EHA are that it allows both time-fixed and time-varying factors into the same model and that it takes care of right-censoring in the data. Right censoring means that for some people the event (withdrawal) may not have occurred yet by the end of measurement (i.e., after 20 rounds). EHA enables estimating transition times despite this information being 'missing' in the dataset (Allison, 2010).

The model estimates the hazard ratio $h_{i,t}$, which is the conditional probability that individual i will drop out in time period t given that they did not do so prior to time point t, i.e. $h_{i,t} = P(T_i = t \mid T_i \geq t)$. For the 20 discrete decision rounds in our study, the hazard ratio is a function of the number of withdrawals relative to the number of subjects at risk of withdrawal. The goal of the analyses is to disentangle what motivates people to participate in the risk-sharing group or to opt out. Subjects of all 43 groups were included if they answered all survey questions included in the analyses (which was true for 423 of 430 subjects). For 216 of them an event was observed (meaning they dropped out at some point during the 20 rounds of the experiment) and because EHA excludes all observations for any round after this decision this brings the total number of observations on the subjects $\times$ rounds level to 5625.

4.4 Results

Figure 2 shows for each of the 21 treatment pairs the number of members of the risk-sharing group (Y-axis) per round (X-axis). The grey bars indicate the predefined random determination of sick people per round. A comparison of the different graphs shows that there is no clear pattern between the treatment conditions. While the theory predicts the HG-L line to end above the HG-H line, this is only the case in about half of the graphs. Moreover, for many pairs we see

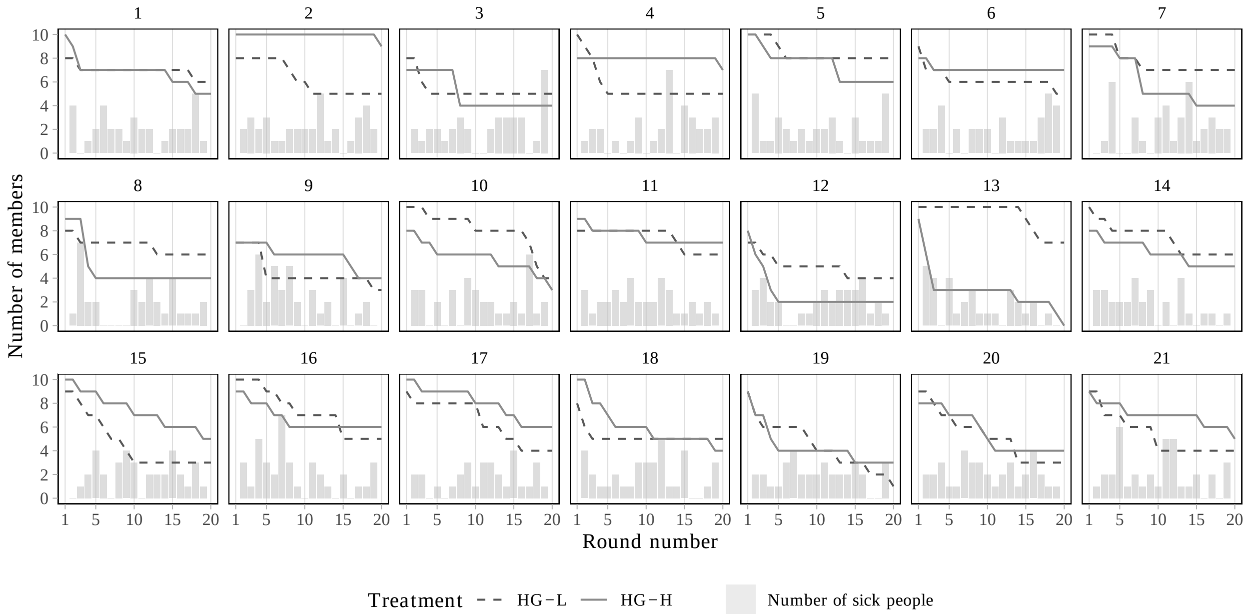


Figure 2. Number of members per treatment per round over all 21 group classifications

the lines crossing each other once or even two or three times. There is no clear trend where one treatment has more members than the other in all rounds. A paired samples t-test for every group $\times$ round combination confirms this. The average number of members in treatment HG-L is $M = 6.35$ ($SD = 1.93$), the average for HG-H is $M = 6.23$ ($SD = 2.12$). The mean difference ($M_d = 0.12$) is not significant ($t(df = 419) = 1.028, p = 0.304$).

What does seem to be apparent from the graphs is that a member who opts out of the risk-sharing group is usually followed by one or several others in the next few rounds. This signals that people seem to react to each other and one opt-out decision could set in motion a cascading event. With the exception of some end-game effects towards round 20, it also seems that when such a withdrawal chain comes to an end, the risk-sharing group remains stable afterwards.

Table 4 reports the results of the Event History Analyses. As can be seen from the left panel of Figure 3, the survival probability is 0.88 for the first round, meaning that 12% of the participants are predicted to never participate at all. After 20 rounds, the survival probability is 0.48. In other words, about half of the subjects have dropped out after 20 rounds (corresponding to 216 events). The second panel of Figure 3 shows the predicted survival plot after all main effects of the predictor variables are included (Model M1a in Table 4). The huge decrease after round 1 suggests that the effects of the variables may not be constant for all rounds, which also treatens the validity of the proportional hazard assumption underlying EHA models (i.e., that ratio of the hazards for any two individuals is constant over time; Allison, 2010). A test of this

Table 4. EHA on the likelihood of withdrawal (216 events, $N = 5625$)

	(M1a)		(M1b)		(M2)	
Received benefit	0.246	(0.184)	0.249	(0.184)	0.249	(0.184)
Risk aversion	-0.015	(0.011)	0.006	(0.015)	0.005	(0.019)
Risk aversion $\times$ t			-0.003^{*}	(0.002)	-0.003^{*}	(0.002)
Risk aversion $\times$ HG-H					0.001	(0.021)
Solidarity	-0.031^{***}	(0.008)	-0.031^{***}	(0.008)	-0.028^{*}	(0.011)
Solidarity $\times$ HG-H					-0.007	(0.016)
Risk	-2.264^{*}	(1.111)	-2.301^{*}	(1.113)	-2.307^{*}	(1.114)
Gender: Female	-0.390^{*}	(0.154)	-0.369^{*}	(0.155)	-0.381^{*}	(0.157)
Gender: Other (not disclosed)	-0.958	(1.014)	-0.929	(1.014)	-0.915	(1.015)
Age	-0.055	(0.048)	-0.055	(0.048)	-0.055	(0.048)
Game theory	-0.495	(0.285)	-0.500	(0.285)	-0.495	(0.286)
Treatment: HG-H	-0.017	(0.140)	-0.025	(0.140)	-0.039	(0.144)
Total members $t-1$	-0.101	(0.067)	-0.105	(0.067)	-0.103	(0.067)
Total left $t-1$	0.452^{**}	(0.162)	0.448^{**}	(0.162)	0.449^{**}	(0.162)
Num. sick members RSG $t-1$	0.392^{***}	(0.060)	0.394^{***}	(0.060)	0.393^{***}	(0.060)
LR χ^2 Test	89.993^{***}	(df = 12)	94.092^{***}	(df = 13)	94.296^{***}	(df = 15)
LR χ^2 Difference Test			4.099^{*}	(df = 1)	0.204	(df = 2)

Note: $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$

assumption indicates that while the overall model does meet this assumption ($\chi^2(13) = 11.8352$, $p = 0.541$) it is violated for the covariate risk aversion ($\chi^2(1) = 5.2161$, $p = 0.022$). We therefore acknowledge this time dependency by adding the interaction between round 1 and risk aversion to the model (Model M1b).

In Model M1b, like in the group-level t-test, the treatment condition does not significantly affect withdrawal rates. Hence, other than for the Broodfonds sample, Hypothesis 1 is not supported. Interestingly, the main effect of risk aversion (which is the effect of risk aversion in round 1) is not significant, but we do find a significant effect of risk aversion $\times$ t. This means that risk averse subjects were indifferent with respect to the decision to join initially, but those risk averse individuals who did join were more likely to remain a member. Hence, we find partial support for Hypothesis 2.

The results do resemble those for the Broodfonds sample with respect to solidarity (Hypothesis 3, supported). The hazard rate for solidarity is significant and negative, which means that people with higher solidarity motives are at lower risk for dropping out. One degree increase on the SVO angle decreases the withdrawal hazard rate by 3%. However, and similar to the Broodfonds sample, the effect of solidarity does not vary depending on the treatment group. Model M2 reports the results of the interactions between the risk heterogeneity treatments and risk aversion and solidarity. The χ^2 difference between Models M1b and M2 did not improve significantly, nor do any of the interaction terms significantly predict withdrawal. Hypothesis 4

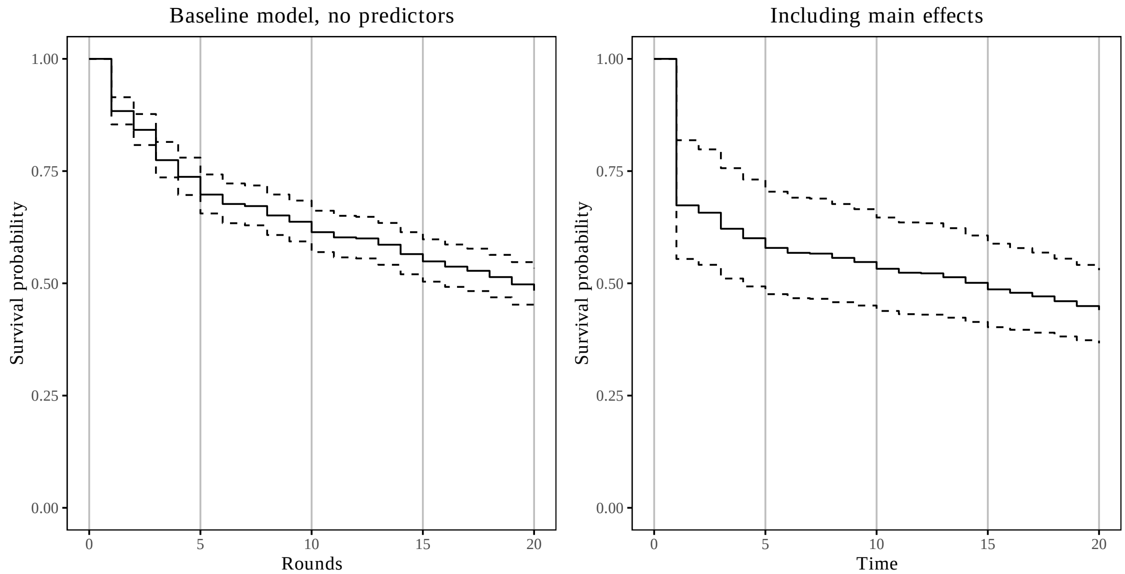


Figure 3. Average survival probability for baseline model and model with predictors

is not supported.

With respect to Hypothesis 5, we take as indicators for interdependencies in decision-making whether subjects respond to other group members getting sick and other group members withdrawing. Both significantly and strongly influence the withdrawal hazard rates. That is, the hazard ratio of withdrawal increases by 48% if one group member was sick and by 56% if one person leaves. Interestingly, the total number of members does not matter. As long as the number stays stable, it does not matter how many people participate. All in all, this supports Hypothesis 5.

Finally, we find that subjects with higher risks are less likely to withdraw. Whether or not the subject received a benefit had, other than for Broodfonds members, no effect on their decision to continue participation. As for our controls, women are less likely to withdraw than men. Age and whether the subject had experience with game theory do not affect withdrawal.

Like for Study 1, we also performed the analyses for the subsample of low-risk subjects in the experiment. As low-risk subjects we selected all subjects with a risk probability below the initial population-level average of $\bar{p} = 0.2$. The results of these analyses are broadly similar to that of the full sample, apart from some effects failing to reach significance under the smaller sample size (i.e., that of risk, gender, and the number of members that left). As a second robustness check, we repeated the analyses in a multilevel model with subjects nested in risk-sharing groups. The reason that we did not report the multilevel event history analyses to begin with is that these do not allow for a test of the proportional hazard assumption. Since the random intercept on the

Table 5. Overview of hypotheses and results

Hypothesis	Exp. direction	Study 1: Broodfonds	Study 2: experiment
H1: Risk heterogeneity	−	−	0
H2: Risk aversion	+	+/0	+/0
H3: Solidarity	+	+	+
H4a: Interaction risk aversion	−	0	0
H4b: Interaction solidarity	−	0	0
H5: Cascades	−		−

group level is small ($\sigma^2 = 0.088$), we could report the unclustered EHA results without much loss of information. All results were found to be robust across the two methods. Detailed results for these two robustness analyses are listed in **????** of **??**.

Table 5 summarizes the results of the different hypotheses for the two studies. The negative effect of risk heterogeneity was found only for the Broodfonds sample. Risk aversion did not directly affect participation in either study, but we found a partial effect for low-risk members in the Broodfonds sample and for all decisions after round 1 in the experiment. Solidarity was positively related to participation in both studies. There are no interactions between heterogeneity and risk aversion or solidarity, but we did find very strong cascade effects in the experiment.

5 Conclusion and Discussion

In this chapter, we aimed to increase our understanding of the motivations and dynamics underlying participation in risk-sharing groups. Driven by the discrepancy between the substantial risk heterogeneity observed in risk-sharing groups in practice and the focus on risk homogeneity in many theoretical discussions, we aimed to further advance theoretical and empirical understanding of risk heterogeneity. How much risk heterogeneity is accepted by the low-risk members before deciding to drop out? How do people deal with the uncertainty introduced by unknown risk differences? To what extent can risk aversion and solidarity explain individual behavior? We hypothesized on the relation between heterogeneity, risk aversion, and solidarity in a dynamic, interdependent setting. These hypotheses were tested with two data sources: a survey among 5192 members of 230 Dutch Broodfonds groups (in which self-employed people share the risk of long-term illness) and an online experiment with 430 British subjects of the Prolific platform. After all, only by combining different methods can we get a better grip onto the (combinations of) factors that are robust in explaining participation in various contexts (Poteete, Janssen, & Ostrom, 2010).

The two data sources rely on vastly different samples and methods. We compared socially and institutionally embedded risk-sharing groups in the real world to general risk-sharing motivations in an artificial setting, and cross-sectional data about participatory intentions to repeated decisions in a controlled environment. Still, the results of the two studies are similar in many regards. The main difference relates to risk heterogeneity, which we found (as hypothesized) to negatively affect commitment in Broodfonds groups, while the experimental manipulation did not result in differences in participation rates in between high and low heterogeneity groups.

This may partly be the result of differences in measurement. In the experiment, risk levels were exogenously determined, but subjects were not informed about the distribution of risk. While we expected them to infer this inequality from observed differences in support requests over time, this may have been too complex. We do see, for instance, a strong reaction to the number of sick group members in the previous round, which could reflect the subject's perception of other group members' risk. Alternatively, while the Broodfonds survey data represents a cross-sectional snapshot, it does capture the effect of risk heterogeneity after a longer time period (ranging from 0 to 11 years), which means the perceptions of group-level risk have had more time to develop, leading to more marked group differences.

With respect to our motivations for participation, we found that the role of risk aversion (central in all risk-sharing literature) is not that clear-cut. Neither for the Broodfonds members nor for the subjects in the experiment did we find a general positive effect of risk aversion such that risk averse members are in general more likely to participate and continue to participate in risk-sharing groups. For the Broodfonds groups, we did find a positive effect if we restricted our sample to only those members whose risk was below the group's average—that is, those members who, strictly speaking, likely pay more to help others than they will receive support in return. For the subjects of the experiment, we found that risk aversion only affected this decision to participate for all rounds after round 1. That is, it increased the likelihood to *continue* to participate, but not the likelihood to join in the first place.

The lack of a general effect for risk aversion could mean several things. Perhaps we did not measure what we intended to measure. This seems unlikely, though, as the instrument has been validated (Falk et al., 2016). Alternatively, it is possible that motivation to share risk is driven not by risk but by loss aversion (Kahneman & Tversky, 1979). However, Vogt (2007) did not find any evidence for loss aversion in a risk-sharing set-up. Hence, we are more inclined to believe that for risk-sharing contexts the evaluation of uncertainty works differently. Particularly in the beginning, without prior knowledge about the other group members or the risk-sharing institution, the risk-sharing group may not successfully reduce uncertainty. For instance, people do not know how many others will participate, whether they continue to participate, how much support they will need, how much participation will cost them, whether the fund will be sufficient to support them, or if they even need support at all.

With so many additional sources of uncertainty, it is not evident which scenario (participating or not participating) is more attractive to a risk averse person. This has been suggested by other scholars observing the low participation rates in mutual insurance groups in low-income countries as well (Platteau et al., 2017) and would require models of ambiguity aversion instead (Elabed & Carter, 2015). Subsequently, for the (low-risk) people who did take the risk to join and built trust in the established cooperation, it becomes possible to disentangle the effect of risk aversion. They experienced that the risk-sharing group can take away a significant share of uncertainty—making it the preferred alternative for risk averse people. In the Broodfonds groups, then, the fact that we only found an effect for low-risk members might be because risk was measured as a self-perception. Those who had to rely on support in the past probably perceive their risk as high(er), so regardless of their (low) risk aversion in general they will be committed to remain a member in the future.

Solidarity was a strong predictor of participation both in Broodfonds groups and for the experimental subjects, albeit only as a main effect, not in interaction with heterogeneity. The strong effect of solidarity is interesting, because it was observed both in the Broodfonds groups, where members know each other, talk to each other, and are embedded within a trusted organizational framework, and in the experiment, where members shared risk with anonymous strangers. Hence, for the experiment, at least, what we measured were altruistic motives in general. For the Broodfonds groups, while we cannot distinguish between general altruism and in-group solidarity (cf. Baldassarri, 2015), it seems safe to say that both are likely to play a role. One indicator, for instance, is that while the number of people requesting support negatively affected participation rates in the experiment, it did not matter for Broodfonds members. This suggests that Broodfonds members, knowing the people who request support, do not mind paying these costs, because they also see how it benefits the recipients. On the other hand, the effect of solidarity in the Broodfonds analyses probably mostly reflects general altruism, because we also controlled for various aspects of social embeddedness (i.e., the number of strong ties and network structures) that probably capture the in-group solidarity aspects.

Finally, we found strong support for cascade effects in the experiment. The more group members got sick, the more likely the subject was to withdraw in the next round. And the more group members withdrew, the higher the chance that others would do so as well. Hence, a sudden increase in the number of sick group members is dangerous, for it often led to withdrawal cascades that could easily lead to overall group-level failure. Does that mean Broodfonds groups and other risk-sharing groups should fear for their resilience? Not necessarily. The effects found in the experiment are magnified compared to a real-world setting. First of all, it was very easy—and even sensible—for participants in the experiment to try out the arrangement for a couple of rounds to get to know the group members and their behavior before deciding to join. Normally, such considerations are made before joining the group. Most people who dropped out in the first few rounds would probably have hesitated to participate in a real-life risk-sharing group and would have ultimately decided against it before the group even started.

Second, the cascades emerged in response to the number of people needing the insurance. This number was inflated to make sure that enough events could be observed to disentangle mechanisms for participation. The average risk of all group members in the experiment was 20% meaning that, while it varied per round, on average two members were expected to be sick in every 10-person population each round. In the Broodfonds group, the average number of benefits was 5% among the respondents, but we know from communication with the organization that the real number is lower. This means that the reflection on whether or not to continue in this group when some people keep reporting ill will be much much slower in real risk-sharing groups. Not only because a 'round' lasts a month rather than a minute, but also because there will be many rounds in which not much happens. Still, despite its inflation in the experiment, risk-sharing groups should actively invest in measures that counter this potential instability.

Before discussing wider implications of our findings, however, some limitations should be taken into account. In the survey among Broodfonds members, first, we relied on a proxy for future cooperation. While commitment is generally considered a proper predictor for future participation (Hauert, 2002; Kollock, 1994), it can only be insofar as future circumstances resemble the present. There is no guarantee that current commitment can be extended to other, more extreme circumstances not part of the current measurement (e.g., a sudden, extreme increase in the number of support requests). Hence, in terms of validity, while the results help to understand the dynamics of risk-sharing under stable risk probabilities, any speculations regarding the effects of changes in risks (and, as a result, risk heterogeneity) remain speculative.

For the experiment, it is important to acknowledge that while experiments generally benefit from greater control, this is more difficult to establish in an online setting. Two possible problems are subjects that carelessly make decisions without properly reading the instructions or assessing the specific decision-making situations, and subjects dropping out before the end of the experiment (Arechar, Gächter, & Molleman, 2018). We took several measures to minimize these problems (such as making sure that unmotivated or impatient subjects had already dropped out before the actual risk-sharing treatment began and warning subjects that low-quality responses were a reason to withhold payment), but some disinterest or dropout could not be avoided.

While each method suffers from some other limitations, the benefit of our multi-method approach is that limitations of one study are mostly not a limitation for the other. For instance, the experiment introduces variation in participation rates and the Broodfonds survey allowed to test risk-sharing dynamics in an applied context. By using two different methods and subject populations, our consistent findings can be considered more robust (Buskens, 2014). There are, however, some aspects missing in both designs that could be addressed in future studies. For instance, while we studied several motivations for participation, we did not take into account issues relating to moral hazard. There is no information available about potential misuse of the common fund in Broodfonds groups, and misuse was not an option for the experimental subjects. This is, however, an important determinant for people's willingness to participate (Van Leeuwen, 2016; Vriens et al., 2019).

Moreover, we lack insights into the role of endogenous group formation. For Broodfonds groups we only know who are currently members of the group. We know nothing about the people who may have considered it but did not join in the end. For the experiment people did not have a choice to form a group endogenously. They could share risk within the assigned population of 10 people, or they could not. People who did not participate in this specific group might have done so in another group, but in our experiment it cannot be disentangled whether people were not interested anyway or were not interested in this specific group. Earlier studies do show that endogenous group formation increases the motivation to share risk for those who manage to create a group (Attanasio et al., 2012; Hegselmann & Flache, 1998). Future research should look at the effects of endogenous group formation both within groups and across groups (for it does introduce a danger of social exclusion) and how this influences the group's risk heterogeneity.

Finally, future research should build on our mixed findings for the role of risk aversion in decision situations characterized by more than one source of uncertainty. In a way, our experiment measured a highly extreme scenario of uncertainty, as people were invited to join an anonymous risk-sharing group. They had to start cooperation from scratch and could not trust in an existing institution. It would be interesting to study emerging risk-sharing groups in the field (e.g., newly established Broodfonds groups) to study whether the perceived risks of participation change over time—or whether the new groups could even build on trust in the general institutional framework. Experimentally, studies should investigate whether collective action in such scenarios with two sources of uncertainty are more likely to be successful—and to attract risk averse people—if the social uncertainty can be taken away (Wit & Wilke, 1998).

Notwithstanding these limitations and open questions for future research, our in-depth approach to studying participation dynamics has generated some important insights. Risk heterogeneity and fluctuations in support requests can be dangerous for future participation levels, so risk-sharing groups should invest to increase commitment levels to such a degree that they can cope with sudden, temporary increases in support requests. Since solidarity proved to be such an important individual driver for participation, risk-sharing groups should actively invest in increasing (or at least maintaining) solidarity levels such that they can be consolidated into general group-level norms about unconditional helping behavior. Moreover, providing information about the general success and the temporary, exceptional nature of cost fluctuations might help to smooth ideas about the group's average risk levels.

References

Akerlof, G. A. (1970). The Market for "Lemons": Quality Uncertainty and the Market Mechanism. *The Quarterly Journal of Economics*, *84*(3), 488–500. 5

Allison, P. D. (2010). *Survival Analysis Using SAS: A Practical Guide, Second Edition.* Cary, NC: SAS Institute Inc. 21, 22

Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2006). Elicitation Using Multiple Price List Formats. *Experimental Economics*, *9*(4), 383–405. 10, 18

Arechar, A. A., Gächter, S., & Molleman, L. (2018). Conducting Interactive Experiments Online. *Experimental Economics*, *21*(1), 99–131. 28

Arrow, K. J. (1971). *Essays in the Theory of Risk-Bearing.* Chicago, IL: Markham Pub. Co. 6, 7

Arrow, K. J. (1984). *Individual Choice under Certainty and Uncertainty.* Cambridge, MA: Harvard University Press. 3

Attanasio, O., Barr, A., Cardenas, J. C., Genicot, G., & Meghir, C. (2012). Risk Pooling, Risk Preferences, and Social Networks. *American Economic Journal: Applied Economics*, *4*(2), 134–167. 3, 29

Azrieli, Y., Chambers, C. P., & Healy, P. J. (2018). Incentives in Experiments: A Theoretical Analysis. *Journal of Political Economy*, *126*(4), 1472–1503. 18

Bakker, D. M. (2019). *Cooperation and Social Control: Effects of Preferences, Institutions, and Social Structure.* Groningen: Rijksuniversiteit Groningen. 18

Baldassarri, D. (2015). Cooperative Networks: Altruism, Group Solidarity, Reciprocity, and Sanctioning. *American Journal of Sociology*, *121*(2), 355–395. 3, 7, 27

Beito, D. T. (2000). *From Mutual Aid to the Welfare State: Fraternal Societies and Social Services, 1890-1967.* Chapel Hill, NC: The University of North Carolina Press. 3

Breer, V. V., & Novikov, D. A. (2015). Threshold Models of Mutual Insurance. *Automation and Remote Control*, *76*(5), 897–908. 2, 3, 6

Burnham, K. P., & Anderson, D. R. (2003). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach.* New York: Springer-Verlag. 12

Buskens, V. (2014). *Coöperatie in Context: Experimentele Sociologie 2.0.* Utrecht: Utrecht University. 4, 29

Buskens, V., & Raub, W. (2008). Rational Choice Research on Social Dilemmas: Embeddedness Effects on Trust. In R. P. M. Wittek, T. A. B. Snijders, & V. Nee (Eds.), *The handbook of rational choice social research* (pp. 113–150). New York: Russell Sage. 4

Charness, G., & Genicot, G. (2009). Informal Risk Sharing in an Infinite-Horizon Experiment. *Economic Journal*, *119*(537), 796–825. 16

Chaudhuri, A. (2011). Sustaining Cooperation in Laboratory Public Goods Experiments: A Selective Survey of the Literature. *Experimental Economics*, *14*(1), 47–83. 6

Chen, D. L., Schonger, M., & Wickens, C. (2016). oTree - An Open-Source Platform for Laboratory, Online, and Field Experiments. *Journal of Behavioral and Experimental Finance*, *9*, 88–97. 18

Coate, S., & Ravallion, M. (1993). Reciprocity Without Commitment: Characterization and Performance of Informal Insurance Arrangements. *Journal of Development Economics*, *40*(1), 1–24. 2, 3, 5, 6

Curvers, E. (2019, sep). *Vertrouwen is het belangrijkste ingrediënt van een broodfonds. Maar houdt het op?* Retrieved from https://www.volkskrant.nl/werken/

vertrouwen-is-het-belangrijkste-ingredient-van-een-broodfonds-maar-waar-houdt-het-op{~}b3850dcc/ 2

Delpierre, M., Verheyden, B., & Weynants, S. (2016). Is Informal Risk-Sharing Less Effective for the Poor? Risk Externalities and Moral Hazard in Mutual Insurance. *Journal of Development Economics*, *118*, 282–297. 6

Elabed, G., & Carter, M. R. (2015). Compound-Risk Aversion, Ambiguity and the Willingness to Pay for Microinsurance. *Journal of Economic Behavior & Organization*, *118*, 150–166. 6, 27

Emery, G. N., & Emery, J. C. H. (1999). *A Young Man's Benefit: The Independent Order of Odd Fellows and Sickness Insurance in the United States and Canada, 1860-1929.* Montreal: McGill-Queen's University Press. 2

Epstein, S. R., & Prak, M. (2008). *Guilds, Innovation and the European Economy, 1400-1800.* Cambridge: Cambridge University Press. 2

Fafchamps, M. (1992). Solidarity Networks in Preindustrial Societies: Rational Peasants with a Moral Economy. *Economic Development and Cultural Change*, *41*(1), 147–174. 4, 5, 6

Fafchamps, M., & Lund, S. (2003). Risk-Sharing Networks in Rural Philippines. *Journal of Development Economics*, *71*, 261–287. 3

Falk, A., Becker, A., Dohmen, T. J., Huffman, D., & Sunde, U. (2016). *The Preference Survey Module: A Validated Instrument for Measuring Risk, Time, and Social Preferences.* 10, 18, 26

Fehr, E., & Schmidt, K. (1999). A Theory of Fairness, Competition and Cooperation. *Quarterly journal of Economics*, *114*(3), 817–868. 6

Genicot, G., & Ray, D. (2003). Group Formation in Risk-Sharing Arrangements. *Review of Economic Studies*, *70*(1), 87–113. 5

Gintis, H., Bowles, S., Boyd, R., & Fehr, E. (2005). *Moral Sentiments and Material Interests: The Foundations of Cooperation in Economic Life.* Cambridge, MA: The MIT Press. 3

Granovetter, M. S. (1978). Threshold Models of Collective Behavior. *American Journal of Sociology*, *83*(6), 1420–1443. 8

Harris, B. (2012). *Welfare and Old Age in Europe and North America: The Development of Social Insurance*. London: Pickering & Chatto (Publishers) Ltd. 7

Harrison, G. W., & Elisabet Rutström, E. (2008). Risk Aversion in the Laboratory. *Research in Experimental Economics*, *12*(08), 41–196. 18

Hauert, C. (2002). Volunteering as Red Queen Mechanism for Cooperation in Public Goods Games. *Science*, *296*(5570), 1129–1132. 10, 28

Hegselmann, R., & Flache, A. (1998). Understanding Complex Social Dynamics: A Plea for Cellular Automata Based Modelling. *Journal of Artificial Societies and Social Simulation*, *1*(3), 1–25. 2, 5, 29

Holzmeister, F. (2017). oTree: Ready-Made Apps for Risk Preference Elicitation Methods. *Journal of Behavioral and Experimental Finance*, *16*, 33–38. 19

Jak, S., & Evers, A. (2010). Een Vernieuwd Meetinstrument voor Organizational Commitment. *Gedrag en Organisatie*, *23*(2), 158–171. 10

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, *47*(2), 263–292. 6, 26

Kollock, P. (1994). The Emergence of Exchange Structures: An Experimental Study of Uncertainty, Commitment, and Trust. *American Journal of Sociology*, *100*(2), 313–345. 10, 28

Krackhardt, D. (1987). Cognitive Social Structures. *Social Networks*, *9*(2), 109–134. 12

Laczó, S. (2014). Does Risk Sharing Increase with Risk Aversion and Risk when Commitment is Limited? *Journal of Economic Dynamics and Control*, *46*, 237–251. 6

Lemay-Boucher, P. (2009). Beninese and Ethiopian Informal Insurance Groups: A Comparative Analysis. *Development Policy Review*, *27*(3), 333–347. 3

Levine, D. K. (1998). Modeling Altruism and Spitefulness in Experiments. *Review of Economic Dynamics*, *1*(3), 593–622. 7

Ligon, E., Thomas, J., & Worrall, T. (2002). Mutual Insurance with Limited Commitment: Theory and Evidence from Village Economies. *Review of Economic Studies*, *69*, 209–244. 3, 5

Lin, W., Liu, Y., & Meng, J. (2014). The Crowding-Out Effect of Formal Insurance on Informal Risk Sharing: An Experimental Study. *Games and Economic Behavior*, *86*, 184–211. 7, 16

Lin, W., Meng, J., & Weng, X. (2019). Formal Insurance and Informal Risk Sharing Dynamics. *Journal of Economic Behavior and Organization*. 7

Meyer, J. P., Allen, N. J., & Smith, C. A. (1993). Commitment to Organizations and Occupations: Extension and Test of a Three-Component Conceptualization. *Journal of Applied Psychology*, *78*(4), 538–551. 10

Murphy, R. O., Ackermann, K. A., & Handgraaf, M. J. (2011). Measuring Social Value Orientation. *Judgment and Decision Making*, *6*(8), 771–781. 18, 20

Orbell, J. M., Schwartz-Shea, P., & Simmons, R. T. (1984). Do Cooperators Exit More Readily than Defectors? *The American Political Science Review*, *78*(1), 147–162. 10

Ostrom, E. (1990). *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge: Cambridge University Press. 6

Ostrom, E. (2005). *Understanding Institutional Diversity*. Princeton, NJ: Princeton University Press. 8

Platteau, J.-P. (1997). Mutual Insurance as an Elusive Concept in Traditional Rural Communities. *Journal of Development Studies*, *33*(6), 764–796. 2, 3, 4

Platteau, J.-P., De Bock, O., & Gelade, W. (2017). The Demand for Microinsurance: A Literature Review. *World Development*, *94*, 139–156. 3, 6, 7, 27

Poteete, A. R., Janssen, M. A., & Ostrom, E. (2010). *Working Together: Collective Action, the Commons, and Multiple Methods in Practice*. Princeton, NJ: Princeton University Press. 25

Tausch, F., Potters, J., & Riedl, A. (2014). An Experimental Investigation of Risk Sharing and Adverse Selection. *Journal of Risk and Uncertainty*, *48*(2), 167–186. 5

Van der Lippe, T., Lippényi, Z., Lössbroek, J., Van Breeschoten, L., Van Gerwen, N., & Martens, T. (2016). *European Sustainable Workforce Survey [ESWS]*. Utrecht: Utrecht University. 10

Van Leeuwen, M. H. D. (2016). *Mutual Insurance 1550-2015: From Guild Welfare and Friendly Societies to Contemporary Micro-Insurers*. London: Palgrave Macmillan. 2, 3, 29

Vogt, S. (2007). *Heterogeneity in Social Dilemmas: The Case of Social Support.* Enschede: PrintPapers Ipskamp BV. 26

Vogt, S., & Weesie, J. (2004). Social Support among Heterogeneous Partners. *Analyse & Kritik*, *26*, 398–422. 2, 3, 6

Vogt, S., & Weesie, J. (2006). Social Support among Heterogeneous Partners: An Experimental Test. *Journal of Economic Interaction and Coordination*, *1*, 215–232. 3, 5, 6

Vriens, E., & Buskens, V. (2020). *Risk-Sharing under Heterogeneity: Exploring Participation Patterns in Situations of Incomplete Information.* 5, 7, 8

Vriens, E., Buskens, V., & De Moor, T. (2017). *Broodfonds Survey [dataset and codebook].* Utrecht: Utrecht University. 2, 9

Vriens, E., Buskens, V., & De Moor, T. (2019). Networks and New Mutualism: How Embeddedness Influences Commitment and Trust in Small Mutuals. *Socio-Economic Review*, *0*(0), 1–22. 4, 7, 10, 11, 15, 29

Vriens, E., & De Moor, T. (2020). Mutuals on the Move: Exclusion Processes in the Welfare State and the Rediscovery of Mutualism. *Social Inclusion*, *8*(1), 225–237. 2, 3

Wit, A., & Wilke, H. (1998). Public Good Provision under Environmental and Social Uncertainty. *European Journal of Social Psychology*, *28*, 249–256. 29