

RESEARCH ARTICLE

Testing evolutionary theories of human cooperation via meta-analysis of microfinance repayment

Dugald Foster , Erik Postma, Shakti Lamba and Alex Mesoudi 

Centre for Ecology and Conservation, University of Exeter Cornwall Campus, Penryn, UK

Corresponding author: Alex Mesoudi; Email: a.mesoudi@exeter.ac.uk

Abstract

Explaining how cooperation evolves is a major research programme in the biological and social sciences. In this study, we tested evolutionary theories of human cooperation in a real-world social dilemma: joint liability microfinance, in which groups of borrowers must cooperate to successfully repay a shared loan. We used pre-registered Bayesian multilevel models to estimate meta-analytic associations between loan repayment and proxies of four evolutionary mechanisms proposed to support cooperation: relatedness, reciprocity, partner choice, and punishment. A systematic search of the microfinance literature yielded 73 effect estimates for 11 proxies of evolutionary mechanisms analysed in 11 separate meta-analyses. Punishment-based variables showed the strongest positive meta-analytic associations with loan repayment, with mixed results for other mechanisms. However, estimates varied widely in their certainty, with generally high levels of between-study heterogeneity. Our results provide some evidence for evolutionary mechanisms supporting cooperation in real-world contexts, but also indicate there are non-generalisable findings and/or reproducibility issues in the microfinance literature.

Keywords: cooperation; punishment; microfinance; meta-analysis; social dilemma[†]

Social media summary: Meta-analyses reveal the evolutionary mechanisms that encourage cooperation in microfinance loan repayment groups.

Introduction

Researchers have identified several mechanisms capable of supporting the evolution and maintenance of cooperation in humans and other species (Apicella & Silk, 2019; Fletcher & Michael Doebeli, 2009; Lehmann & Keller, 2006; Nowak, 2006; West et al., 2007a; West et al., 2007b), and which could explain why humans are more or less likely to cooperate under particular circumstances. Theoretical models, laboratory experiments, and field studies have all been used to explore the viability of these mechanisms in promoting cooperation (Apicella & Silk, 2019; Nowak, 2006). However, most studies have tested the effects of individual mechanisms in isolation, with few comparing mechanisms simultaneously using real-world human behavioural data. Due to its global implementation across diverse human ecologies, populations, and cultures, joint liability microfinance provides an opportunity to test multiple evolutionary mechanisms of cooperation, including kin selection, reciprocity, partner choice, and punishment, within the context of the same real-world social dilemma.

Microfinance loans are designed for people who cannot access traditional loans from banks, either because they are unable to provide collateral such as land or property required as a guarantee against non-repayment or because they are excluded from financial services due to geographic or economic barriers (Armendáriz & Morduch, 2010). Some microfinance institutions (MFIs) offer ‘joint liability’

loans (sometimes called ‘group loans’ or ‘solidarity loans’), provided to a group of borrowers who assume collective responsibility for repaying the loan. Once a group has been formed by individuals interested in taking out a loan, the group signs a loan contract with the lending institution stipulating the size of the loan, the interest rate, and the repayment schedule. Crucially, all members of the group must accept the terms of joint liability attached to the loan. Borrowers in the group then receive and invest their loans at their discretion. At agreed intervals, the group convenes at repayment meetings where members repay their individual shares of the loan to the lending institution. If a borrower does not attend the meeting or fails to repay his/her share, the other group members are collectively liable for his/her repayment. If the group fails to repay the loan in full, all members are banned from accessing further loans and could face additional penalties (Armendáriz & Morduch, 2010).

Prominent theoretical models in economics assume that joint liability loans allow MFIs to insure themselves against the risk of individual borrowers failing to repay, by replacing material collateral with social collateral (Armendáriz, 1999; Besley & Coate, 1995; Conning & Udry, 2007; Ghatak & Guinnane, 1999; Stiglitz, 1990; Van Tassel, 1999; Varian, 1990). These models highlight that the shared responsibility of repaying a group loan creates a social dilemma: a scenario in which ‘selfish’ individuals who maximise their own success outperform cooperative individuals who pay a cost for the sake of the group, but where groups of cooperators outperform groups of non-cooperators (Dawes, 1980; Olson, 1971). Hence, individual- and group-level interests conflict. Under joint liability, microfinance borrowers seeking to maximise their immediate financial returns have an incentive to default on their repayment dues, in effect receiving a free loan, and to free-ride on the repayment efforts of others in their loan group. Yet the loan group succeeds only if the total loan amount is repaid, thereby ensuring continued access to future and potentially larger loans for all group members. To successfully repay a group loan, borrowers must therefore cooperate: first by repaying their own share of the loan, and if necessary by enforcing the payment of money owed by any defaulting group members, or else paying it themselves (Besley & Coate, 1995).

Microfinance loans are popular globally, with around 140 million customers borrowing 125 billion US dollars in 2018, and with large markets operating in Southeast Asia, sub-Saharan Africa, Latin America and the Caribbean, and East Asia and the Pacific (Microfinance Barometer, 2019). Given the popularity of microfinance, and because of the mixed evidence for its success in poverty reduction (Meager, 2019; Roodman, 2012), it is important to understand what prevents loan groups from becoming overrun by free-riding defaulters. While many MFIs regularly report high repayment rates for group loans (Cull *et al.*, 2007), there is considerable variation in loan repayment efficacy across time and space (Attanasio *et al.*, 2015; Carpena *et al.*, 2013; Giné & Karlan, 2014; Mahmud, 2019). For example, the Grameen Bank, one of the largest MFIs in Bangladesh, has reported variation in repayment rates for group loans from 100% of loans repaid in one district to 30% in another (Woolcock, 1999). During the Indian microfinance crisis of 2010, national group loan repayment rates dropped from 98% to 10% before subsequently recovering (Haldar & Stiglitz, 2016). We hypothesise that this variation in repayment rates reflects variation in the ability of loan groups to solve the social dilemma underlying the joint liability model. We therefore predict that if joint liability lending relies on cooperation between borrowers, factors that facilitate cooperation between borrowers will show a positive association with group loan repayment efficacy.

Lamba (2014) first identified the potential of using joint liability lending as a model system to test evolutionary theories of cooperation. This was developed by Gehrig *et al.* (2021), who qualitatively summarised the effects of a range of variables representing evolutionary mechanisms of cooperation on loan repayment. Surprisingly, genetic relatedness between group members was found to be largely negatively associated with repayment efficacy, although prior interaction (*i.e.* reciprocity) and partner choice were both more positively associated. As noted by Gehrig *et al.* (2021), however, the vote counting method they used, in which statistically significant findings were tallied across studies, has severe limitations, including an inability to assess the magnitude of effects and a failure to assign differential weights to studies based on characteristics such as study quality and statistical power

(Higgins & Thomas, 2019). Nonetheless, Gehrig et al. (2021) provided a useful preliminary evaluation of the loan repayment literature as a model system to test evolutionary theories of cooperation, which merits further investigation.

Here we conduct quantitative, theoretically informed, and pre-registered meta-analyses of predictors of joint liability loan repayment to estimate average associations between repayment and proxies of evolutionary mechanisms of cooperation and to assess the uncertainty, heterogeneity, and generalisability of these estimates. We find weak and largely inconsistent evidence of associations between our evolutionarily informed predictors and loan repayment, with substantial uncertainty in our estimates predominantly caused by a large degree of between-study heterogeneity that limits the generalisability of our findings. However, in conducting our meta-analyses, we identify several limitations in the existing microfinance literature that generate constraints in terms of causal inference, including inconsistent measurement of effect estimates and predictor and outcome variables, endogeneity in loan group formation, uncontrolled variation in study design and settings, and over-reliance on cross-sectional, non-experimental designs.

Study hypotheses and predictions

We tested four hypotheses based on prominent evolutionary mechanisms that could plausibly support cooperation among microfinance borrowers. These mechanisms all act by increasing assortment between cooperators so that cooperators are more likely to interact with other cooperators, and less so with free-riders (Fletcher & Michael Doebeli, 2009). We originally planned to test three more hypotheses (see study pre-registration at <https://osf.io/9e6ay>), but the microfinance literature we reviewed was unable to provide the necessary data. For each hypothesised mechanism, we make predictions regarding its effect on group loan repayment outcomes. We declare no predictions concerning the relative magnitude of the effects of these mechanisms, which may interact in complex ways.

Kin selection via common ancestry

Loan groups composed of more relatives/more closely related borrowers should exhibit higher repayment rates compared to groups composed of fewer relatives/less closely related borrowers. Theoretical models demonstrate that cooperation can evolve when it is preferentially directed towards genetic relatives who are more closely related to the cooperator due to their common ancestry, relative to the population mean (Hamilton, 1964a, 1964b). Cooperative behaviours that evolve due to their positive impacts on the success of relatives despite their costs to the actor are said to result from 'kin selection' (Grafen, 2007; Maynard Smith, 1964). Given some genetic basis to cooperative behaviour and a human evolutionary history of cooperation driven by kin selection, assortment between relatives could correspond to assortment between cooperators. In joint liability loan groups composed of genetic relatives, the benefits provided to relatives by successful repayment could therefore outweigh the costs of cooperation.

Prior interaction and reciprocity

Loan groups composed of borrowers with experience or knowledge of one another's past cooperative behaviour should exhibit higher repayment rates. Cooperation can evolve when it is preferentially directed towards individuals known to be cooperators based on their past behaviour (Trivers, 1971). Direct reciprocity describes an actor's decision to cooperate with an individual who has previously cooperated with that actor (Axelrod & Hamilton, 1981; Trivers, 1971). Indirect reciprocity describes an actor's decision to cooperate based on their knowledge of prior interactions between an individual and others, or by using information regarding that individual's reputation for cooperation (Nowak & Sigmund, 1998). In the context of microfinance, measures of prior interaction amongst

group members such as whether group members knew each other before forming the group, the length of time the group has been in existence, the frequency of group meetings, and the geographical proximity of group members should all provide information about other group members' past histories of cooperation and therefore affect repayment via reciprocity. For example, group members might choose to help defaulters repay a loan when they themselves have previously been the beneficiary of repayment assistance from that defaulter (direct reciprocity). Or, they may help a defaulter repay a loan because they know that this defaulter has always paid off their loans in the past and this failure to repay is a one-off event rather than a repeated pattern of free-riding (indirect reciprocity). While prior interactions might also facilitate the *failure* to repay a loan if those defaulters are known instead to be free-riders, such instances are likely to be rarer, as failure to repay loans will preclude continued access to subsequent loans. In any case, free-riding is easier when there is a lack of information about past cooperation, hence our prediction that more frequent prior interaction should on average increase repayment rates.

Partner choice

Loan groups formed by borrowers themselves should exhibit higher repayment rates compared to groups formed by a lending institution or groups formed at random. Previous work has shown that cooperation can evolve when cooperators can choose to preferentially interact with other known cooperators and actively dissociate from free-riders, leading to positive assortment (Aktipis, 2004; Hruschka & Henrich, 2006; McNamara *et al.*, 2008; Noë & Hammerstein, 1994). While reciprocity describes decisions to cooperate given the partners available once a group has been formed (e.g. whether to help another group member repay their loan), it does not encompass the choice of available partners at the time of group formation. Selecting cooperative partners from the outset through partner choice may do more to facilitate cooperation than individual decisions made during future interactions (Hruschka & Henrich, 2006). While some loan groups are formed with input from the lending institution (Feigenberg *et al.*, 2013; Karlan, 2007), others are formed by borrowers themselves (Banerjee *et al.*, 2015; Morduch, 1999). Some MFIs also allow loan groups to alter their group membership, for example, by ejecting a persistently defaulting borrower (Carpenter & Williams, 2014), which could also increase assortment between cooperators.

Punishment

Loan groups which carry out sanctions against defaulters, and those that conduct higher levels of monitoring, should exhibit higher repayment rates. Punishment has been identified as central to solving social dilemmas, particularly for large groups of unrelated individuals that have been formed (initially) without assortment (Ågren *et al.*, 2019; Fehr & Gächter, 2002; Frank, 1995). By attaching an extra cost to free-riding, punishment can alter the relative costs and benefits of cooperating, so that cooperation offers the highest pay-offs (Lehmann & Keller, 2006), and would-be free-riders turn to cooperation, thereby increasing assortment. Studies of real-world collective action problems (Herrmann *et al.*, 2008; Kaplan *et al.*, 1985; Ostrom *et al.*, 1992; Wiessner, 2005) and laboratory experiments (Ertan *et al.*, 2009; Fehr & Gächter, 2002; Gülerk *et al.*, 2006; Yamagishi, 1986) support the importance of punishment as a mechanism for maintaining cooperation. While policing free-riders and coordinating punishment may be associated with their own costs (Boyd & Richerson, 1992; Smaldino *et al.*, 2013), cooperative punishers can be rewarded for their efforts through the receipt of additional benefits, for example, via the ostracism of free-riders from the group, resulting in an increased likelihood of interacting with other cooperators (Bowles & Gintis, 2004; Frank, 2003). Similarly, if the costs of punishment are shared between a sufficiently large proportion of group members, punishment is both more effective and the individual costs of delivering punishment decline

(Boyd et al., 2010, 2003). While little is known about the dynamics of actual punishment within joint liability loan groups (Solli et al., 2015), there is anecdotal evidence of groups cooperating to punish defaulters, ranging from mild social shaming of defaulters to collectively confronting them at home, to seizing their household goods and demanding payment for their release (Solli et al., 2015).

Summary of hypotheses

We pre-registered four primary predictions based on evolutionary mechanisms capable of supporting cooperation: relatedness, reciprocity, partner choice, and punishment. We also pre-registered exploratory meta-analyses of three variables that appear regularly in the loan repayment literature (Group size, Borrower age, and Borrower sex) but which do not represent discrete mechanisms or clearly align with any particular hypotheses. We did however have theoretical reasons to investigate these variables, all of which have been implicated in evolutionary theories of cooperation and assortment. For example, some models suggest that cooperative behaviour and reputations are harder to track in larger groups (Suzuki & Akiyama, 2005), and experiments have shown that the reduction in marginal returns with increasing group size can reduce the benefits of cooperation (Isaac & Walker, 1988). Depending on the method of group formation, larger groups could also require a greater proportion of cooperators to maintain cooperation, potentially creating an obstacle to the evolution of cooperation via reciprocity (Boyd & Richerson, 1988). We therefore predicted that larger loan groups would show worse repayment efficacy on average. We did not make directional predictions for borrower sex or age, but still planned to model their average effects, as this could provide insights of applied relevance to microfinance lending. We also tested the role of each of these three variables in explaining heterogeneity between studies.

Methods

We used meta-analysis (Gurevitch et al., 2018; Moreau & Gamble, 2022; Quintana, 2015) to estimate overall associations between a range of variables and loan repayment. To inform our study, we conducted a pilot analysis of the 41 studies included in Gehrig et al. (2021) (see Supplementary Materials SM1). This pilot analysis qualitatively evaluated and compared the statistical models and variable types reported in the 41 studies from Gehrig et al. (2021) in order to develop a feasible statistical meta-analytic strategy (note that no formal meta-analyses were performed in the pilot study itself). Based on the pilot study, we selected 12 variables to meta-analyse (see Table 1). We chose these variables based on their frequent occurrence in group loan repayment studies and their suitability as proxies for the evolutionary mechanisms discussed above. Nine variables are directly linked to specific mechanisms, while three are ambiguous in their relationship to any particular mechanism but may influence the returns from cooperation (Table 1). We pre-registered our predictions and methods on the OSF platform, using the PRISMA-P template (Moher et al., 2015) to specify our methods, available at <https://osf.io/3y9fu>.

Meta-analysis eligibility criteria

Studies had to meet three criteria to be eligible for inclusion in the meta-analyses:

1. *Relevant predictor variables and outcome measure*

Studies had to report quantitative analysis of a predictor variable relevant to 1 of the 12 variables we planned to meta-analyse (Table 1), and its association with an outcome measure of joint liability loan repayment. This includes experimental and cross-sectional studies of group loan repayment and studies comparing repayment between group and individual loans.

Table 1. Common predictor variables analysed in loan repayment studies and included in our meta-analysis, and their association with evolutionary mechanisms of cooperation

Variable	Evolutionary mechanism	Justification
VC1. Relatives in the loan group (binary)	Common ancestry	The presence of family members in loan groups provides an indication of genetic relatedness within groups.
VC2. Prior acquaintance of group members (binary)	Prior interaction	Borrowers who know each other prior to joining a group are more likely to have information about each other's cooperative histories.
VC3. Group tenure (months)	Prior interaction	Borrowers belonging to more long-lasting loan groups are more likely to have information about each other's cooperative histories.
VC4. Frequency of group meetings (weeks)	Prior interaction	Borrowers who meet more frequently are more likely to have information about each other's cooperative histories.
VC5. Geographic proximity of group members (km)	Prior interaction	Borrowers who live or work closer to one another are more likely to have information about each other's cooperative histories.
VC6. Management of group membership (binary)	Partner choice	Loan groups which actively manage their membership (e.g. by screening potential members) enable the selection of cooperative partners.
VC7. Group sanctions (binary)	Punishment	Loan groups which impose their own penalties on defaulters create additional costs for non-repayment.
VC8. Peer monitoring (binary)	Punishment	Borrowers who experience higher levels of monitoring by their loan group are more likely to increase their repayment effort to avoid punishment.
VC9. External monitoring (binary)	Punishment	Borrowers who experience higher levels of monitoring by the loan officer or lending institution are more likely to increase their repayment effort to avoid punishment.
VC10. Group size	N/A	No particular mechanism
VC11. Borrower age (years)	N/A	No particular mechanism
VC12. Borrower sex	N/A	No particular mechanism

Note: 'Binary' indicates that the predictor is of the form 'yes/no' or 'present/absent'; otherwise units are provided for continuous predictors.

2. Publication status

Studies must have undergone peer review or some form of expert assessment, including scientific journal publications, unpublished but examined doctoral theses, and working papers prepared for conferences or journal submission. Here, we aimed to strike a balance between ensuring a high quality of studies and including relevant results, which may exist outside journal publications. Any study that did not meet quality criteria as measured by our risk of bias threshold (see below) was excluded from our main analyses.

3. Years and languages considered

We considered all English-language studies available online before July 2021.

Information sources

For the primary literature search, we used Google Scholar, Web of Science, EconLit, and ProQuest search engines, with the aim of improving search performance by combining results from multiple databases and multiple search engines (Bramer *et al.*, 2017). To check for eligible papers missed in

our primary search, we also hand-searched the reference lists of eligible studies and used the CoCites literature search tool (Janssens et al., 2020) to identify the most frequently co-cited papers of the most relevant publications as defined by our inclusion criteria.

Search strategy

Search terms are presented in Table S1a. In addition to general terms, for each literature search, we included an additional search term for the relevant variable (Table S1b). If any search returned too many results to screen due to resource constraints, only the first 300 search results of each database were screened. This cut-off follows the recommendation of Haddaway et al. (2015) for maximising relevant results from the academic literature in Google Scholar. To check this cut-off, we confirmed in our pilot study that Google Scholar results 300–500 did not contain any eligible studies.

Study records

We used the systematic review management software Rayyan (Ouzzani et al., 2016) to manage study records, screen abstracts, and track inclusion/exclusion justifications. A folder containing all search results (*FosterMA7_SearchResults*) is available from the OSF project page. Data were extracted from studies, coded, and entered into our data extraction template. A research assistant double-coded study effect estimates in order to identify any coding errors. Discrepancies were resolved by consulting the original text and discussion with the research team. As recommended by Muka et al. (2020), we piloted our data extraction form using the 41 loan repayment studies identified in Gehrig et al. (2021), all of which were included in our final list of eligible studies.

To mitigate inclusion bias and HARK-ing (Hypothesising After the Results are Known) (Kerr, 1998), the primary reviewer (DF) made efforts to avoid learning the results of loan repayment studies that would potentially be included in the meta-analysis, only reading the Methods sections and not the Results sections. All hypotheses were derived a priori from evolutionary theories of cooperation, and without reference to the results of group loan repayment studies.

Outcomes and prioritisation

Primary outcome variables included all measures of loan repayment efficacy, including measures of default, repayment delinquency, arrears, and loan group survival time. We interpret greater repayment efficacy as groups displaying higher levels of cooperation. This may be intentional on the part of the borrowers, where non-repayment represents an intentional attempt to free-ride on the efforts of other group members. In other cases, borrowers may fail to repay their loan due to bad luck or circumstances outside the borrower's control, without an explicit intention to free-ride. However, even when non-repayment is not intentional, risky loan investment decisions by borrowers and a lack of sufficient monitoring by their loan group members are themselves plausibly demonstrative of a lack of cooperation in the loan group (Ghatak & Guinnane, 1999). Ultimately, borrowers who fail to repay their share of a group loan prioritise their individual welfare over that of the loan group, so that lower repayment efficacy can generally be considered indicative of non-cooperation within the loan group.

Risk of bias in individual studies

We were unable to find a tool that suited our need to assess the risk of bias for observational studies in non-clinical settings. We therefore designed a new tool – ROBOS ('Risk of Bias in Observational Studies'), which requires answering prompts to record factual information about how studies were conducted. This information is used to form judgements of the risk of bias for each study across four domains of potential bias (study design, sampling method, data collection, and analysis methods), and for the study overall. For our primary analysis, we restricted our data estimates to those from

studies scoring a ‘Low’ or ‘Moderate’ risk of bias, and conducted sensitivity analyses to test the effects of including results from studies with a ‘High’ risk of bias. ROBOS assessments conducted for all studies are available at <https://osf.io/wsdjn>.

Data synthesis

Effect estimates are reported using a variety of statistical measures across different studies and the best way to compare or combine them in a meta-analysis is non-trivial. Our pilot study yielded 14 different effect measures reported across loan repayment studies, the majority of which were log odds. Thus, based on our pilot analyses (see pre-registration and SM1), where we explored multiple ways of comparing and combining different effect measures while maintaining viable sample sizes and interpretability of the meta-analytic estimate, we decided to conduct meta-analyses of log odds outcomes only. An additional advantage of analysing log odds is that its posterior distribution is close to normality, even for relatively small sample sizes (Gelman et al., 2013). To enable comparison between effect estimates from studies with different but comparable outcome variables, we coded all estimates so that positive effect signs reflect increased loan repayment (cooperation), and negative effect signs reflect reduced loan repayment (non-cooperation).

Meta-analytic model

We expected that many of the eligible studies in our sample would not meet our strict risk of bias threshold, resulting in small sample sizes for our primary analyses. For small sample sizes, Bayesian methods are often preferable due to improved estimation relative to frequentist alternatives (Meager, 2019; Williams et al., 2018) and for avoiding the common problems associated with null hypothesis significance testing (Gelman, 2018). We therefore fit Bayesian multilevel models to estimate meta-analytic associations for each of our 12 variables on group loan repayment. We used a random effects meta-analytic model in which effect sizes are nested in studies, and we assume that the ‘true’ effect of the predictor being modelled differs between studies, either due to actual variation in the effect, or due to sampling variability. The Bayesian meta-analytic model takes the form:

$$\begin{aligned} y_i &\sim \text{Normal}(\theta_i, \sigma_i) \\ \theta_i &\sim \text{Normal}(\mu, \tau) \\ \mu &\sim \text{Normal}(0, 1) \\ \tau &\sim \text{HalfCauchy}(0, 1) \end{aligned}$$

where the effect estimate y from study i is assumed to be drawn from a Normal distribution with mean θ_i and standard deviation (sd) σ_i . We assume that θ_i is itself drawn from a Normal distribution, where μ is the mean of the distribution of effect sizes, and τ is the sd around that mean. These two parameters represent the two primary estimands of any meta-analysis: μ is the average association across all exchangeable studies, and τ is the sd of effects between studies. A further advantage of Bayesian models is the ability to interpret posterior distributions as probability distributions of the effects of interest (Salanti et al., 2019). Our models provide us with posterior distributions for the meta-analytic effect size μ , and for the heterogeneity of the underlying effect τ (the between-study variance). This allows us to model the uncertainty associated with the meta-analytic effect for each of our predictors of loan repayment, and τ additionally allows us to estimate how likely it is that our meta-analytic effects are generalisable to other settings (Vivalt, 2020).

Because individual studies show conflicting results regarding the direction and magnitude of effects (Gehrig et al., 2021), we have little robust information to inform our choice of priors. For our meta-analyses of log odds estimates, we therefore assigned a Normal prior centred on zero for μ ,

to spread the probability diffusely and allow the data to dominate our inferences, as recommended by Gelman (2019). For τ , we used a half-Cauchy prior, a distribution with a broad peak around zero and a heavy tail, recommended as a generally applicable prior without strong accompanying assumptions (Gelman, 2006; Polson & Scott, 2012), and given that the between-study variance must be positive (Brockwell & Gordon, 2001). This choice of prior also reduces the risk of inflated estimates for the summary effect μ (Williams et al., 2018). We explored the impact of our choice of priors for both μ and τ via sensitivity analyses (Table S3).

The effect estimates that provide data for the model are weighted using the generic inverse variance method, in which each study's effect estimate is given a weight equal to the inverse of its variance. This method provides more weight for studies with larger samples and therefore estimates with smaller standard errors, and thus reduces the imprecision of the meta-analytic effect estimate (Higgins & Thomas, 2019). We use the median of the posterior distribution as point estimates, as this parameter value minimises expected absolute loss (McElreath, 2016), and reduces the weight attached to outliers. Generalisability was assessed using between-study variance, τ , with higher between-study variance indicating lower generalisability of meta-analytic estimates (Supplementary Materials SM6). Following Spiegelhalter et al. (2004), values of $0.1 < \tau < 0.5$ were interpreted as showing 'reasonable generalisability', $0.5 < \tau < 1$ as 'low generalisability', and $\tau > 1$ as 'very low generalisability'. Publication bias was assessed using Bayesian Copas selection models (Copas, 1999; Copas & Shi, 2001), which assume that the probability of a study being selected (i.e. published) is a function of the size of its effect estimate and the standard error of this estimate (Supplementary Materials SM8). For each variable, we calculated Bai's D (Bai et al., 2020), for which values close to 0 indicate that posterior distributions of our original model and the equivalent Copas selection model are nearly identical (suggesting negligible publication bias), and values close to 1 indicate non-overlapping distributions (suggesting severe publication bias).

We used R (R Core Team, 2020) via RStudio (RStudio Team, 2020) and *brms* (Bürkner, 2017) to fit our Bayesian meta-analytic models in the Stan programming language (Carpenter et al., 2017). Stan enables Markov Chain Monte Carlo sampling with the No U-Turn Sampler (Hoffman & Gelman, 2014; McElreath, 2016) to approximate the posterior distributions of model parameters. The analysis scripts we used are available from: <https://github.com/dugaldfoster/MicrofinanceMetaAnalysis>.

Results

Literature searches

Our 12 separate literature searches returned 3,349 studies. Screening of these studies revealed 225 studies featuring quantitative analysis of the relationship between joint liability loan repayment and at least one of our predictors of interest. From these studies, we extracted 245 relevant effect estimates, as some studies contained multiple predictors of interest. Together with the 85 estimates obtained in our pilot study, this provided a total of 330 estimates (Table 2). In total, 114 of these effect estimates were found to be comparable and 74 of them had a Low/Moderate risk of bias. Table S2 describes literature search results in greater detail.

Descriptive statistics

Our dataset of 330 eligible studies from which estimates were drawn spans 36 countries (Figure 1). The main regions represented are North and Latin America, sub-Saharan Africa, and South Asia, reflecting the areas in which group lending is most prevalent. The greatest number of microfinance repayment studies in our dataset are from Ethiopia (seven studies), India (six studies), and Bangladesh (five studies). Moreover, 63% of borrowers across all studies in our dataset were female. The mean age of borrowers was 36.52 years (sd 7.60), and the mean number of borrowers per loan group was 7.16 (sd 4.08). Study sample sizes ranged from 38 to 81,852 borrowers, with a median of 240 (sd 15,701).

Table 2. The number of relevant effect estimates identified in our pilot study (column ‘Pilot study’), main study (column ‘Main study’), the total number of estimates (column ‘Total’), the number of exchangeable or comparable estimates, i.e. estimates from studies using the same predictor measures, outcome measures, and effect measures (column ‘Comparable’), and the number of comparable studies rated as having a low or moderate risk of bias (column ‘Comparable and low/moderate RoB’)

Variable	Pilot study	Main study	Total	Comparable (% of total)	Comparable and low/moderate RoB (% of total)
VC1. Relatives in the loan group	10	8	18	2 (11)	2 (11)
VC2. Prior acquaintance of group members	18	11	29	8 (28)	6 (21)
VC3. Group tenure	17	32	49	12 (24)	7 (14)
VC4. Frequency of group meetings	5	10	15	2 (13)	0 (0)
VC5. Geographic proximity of group members	11	14	25	7 (28)	5 (20)
VC6. Member management	NA	21	21	12 (57)	7 (33)
VC7. Group sanctions	15	17	32	10 (31)	6 (19)
VC8. Peer monitoring	9	19	28	9 (32)	7 (25)
VC9. External monitoring	NA	16	16	8 (50)	3 (19)
VC10. Group size	NA	31	31	15 (48)	9 (29)
VC11. Borrower age	NA	34	34	13 (38)	11 (32)
VC12. Borrower sex	NA	32	32	16 (50)	11 (34)
Total	85	245	330	114 (35)	74 (22)

Risk of bias assessments

Our dataset contains estimates from 3 studies rated as having a low risk of bias, 31 studies rated as having a moderate risk of bias, and 14 studies rated as having a high risk of bias. As specified in the study pre-registration, our primary analyses exclude estimates from studies rated as having a high risk of bias. Results of the sensitivity analyses testing the impact of including high-risk-of-bias studies are reported in Table S4.

Meta-analysis results summary

We collected sufficient data to meta-analyse the association between 11 variables and loan repayment, with insufficient data for VC4 Frequency of Group Meetings. All models displayed suitable convergence ($R\text{-hat}$ values all < 1.01) and stable trace plots (see McElreath, 2016). Our main results are displayed as a forest plot of odds ratio (OR) estimates in Figure 2. Figures S1–S11 show separate forest plots for each variable, showing OR estimates for each study that contributed to that variable’s overall estimate. See Table S4 for sensitivity analyses conducted separately for each variable exploring the impact of (a) using alternative priors, (b) splitting studies by level of analysis (repayment measured at the individual- vs group level), and (c) including studies with a high risk of bias. Our results are generally robust to alternative modelling choices, with minor differences in results when using alternative

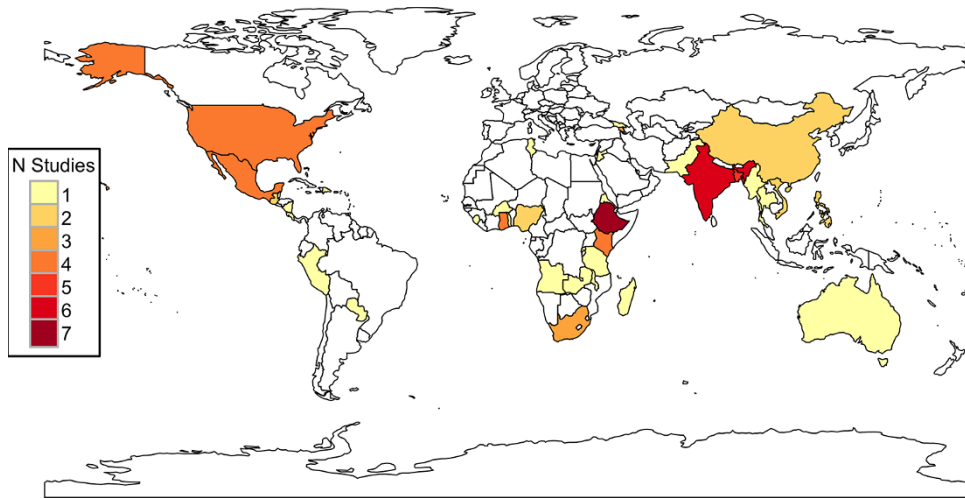


Figure 1. A map of the distribution and frequency of occurrence of the 36 countries from which studies of joint liability loan repayment have been included in our dataset. Many more loan repayment studies exist than are shown here, as our data are restricted to studies that include quantitative analysis of at least 1 of our 12 predictors of interest.

priors, and results shifting slightly and unsystematically when including high-risk-of-bias studies. For some variables, there is evidence of differences between meta-analytic estimates of studies measuring repayment outcomes at the individual vs the group level, although 95% credible intervals (CI) in all sensitivity analyses include OR of 1 indicating no certain effect. Datasets used in the final analysis are available from the OSF project page (<https://osf.io/wsdjn/>) and the GitHub page (<https://github.com/dugaldfoster/MicrofinanceMetaAnalysis>) to facilitate comparison and to provide a resource for future systematic reviews and meta-analyses.

The 95% CI for all 11 meta-analytic estimates include OR of 1 (Figure 2), so we cannot confidently rule out null associations for any variable. We can however use the empirical cumulative distribution function to calculate the probability that any OR estimate is greater or less than 1, indicating a positive or negative association with loan repayment, respectively. In order of probability, the following variables are most likely to have positive associations with group loan repayment: peer monitoring between borrowers (94%), group sanctions (90%), larger groups (90%), female borrowers (87%), and older borrowers (86%). These associations vary widely in their magnitudes, showing between a median 2% increase in the odds of repayment for borrowers who are a year older, and a median 89% increase in the odds of repayment for loan groups that implement sanctions for non-repayment. These estimates also vary widely in their certainty, with 95% CI spanning a difference in OR of 0.1 for the association between borrower age and repayment, and 3.77 for group sanctions. The generally wide CI make it difficult to distinguish clearly between likely negative, null, or positive associations, although the ability to characterise this uncertainty is an advantage of our Bayesian approach.

Two variables are more likely to have negative associations with successful group loan repayment: external monitoring of the group (87%) and closer geographic proximity between members (75%). Magnitudes range from a 7% decrease in the odds of repayment for groups with geographically closer borrowers, and a 42% decrease in the odds of repayment for groups that are externally monitored.

There is a lower probability of positive associations with repayment for acquaintance between borrowers prior to joining a loan group (65%), and management of membership by the group through self-selection and screening (54%). There is also a lower probability of negative associations with repayment for a larger proportion of relatives in a loan group (65%), and longer group tenure (56%).

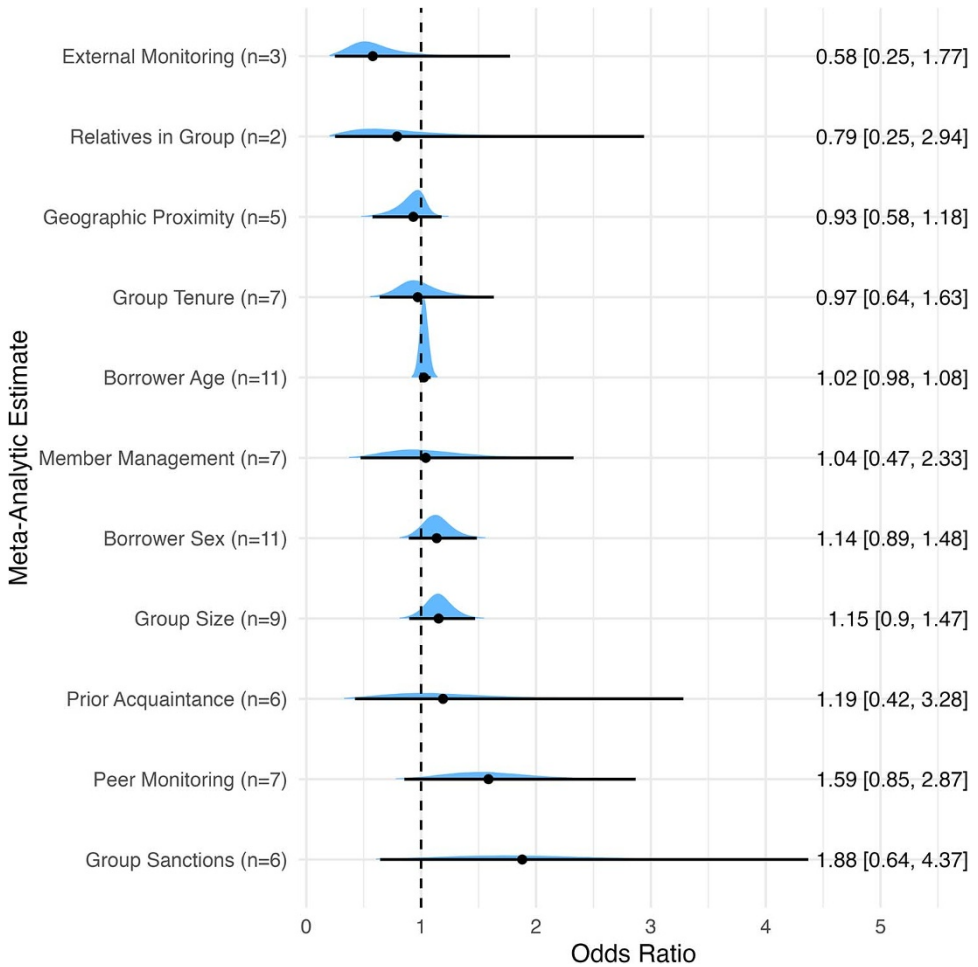


Figure 2. Posterior distributions (blue shading), posterior medians (black dots), and 95% credible intervals (black solid lines) of odds ratios for the meta-analytic association between each variable and group loan repayment, arranged by magnitude of association. If the posterior distribution includes the dotted line representing an odds ratio of 1, then we cannot rule out a null effect. Number of studies for each variable, *n*, is shown in brackets.

These four variables are likely to have either weak or null associations with loan repayment, although the posterior distributions for some estimates have long tails.

There were an insufficient number of comparable estimates of the association between loan group meeting frequency and loan repayment to meet our criteria for meta-analysis. These studies are characterised by a general lack of consistency in measures of meeting frequency, making it difficult to assess its overall relationship with loan repayment. Studies used variables as diverse as ‘1–5 indicators, with higher value indicating more frequent meetings’ (Bastelaer & Leathers, 2006), ‘group meetings regularly attended’ (Singh & Padhi, 2017), and the treatment effect for a randomised control trial (RCT) that manipulated meeting frequency for a treatment group from fortnightly to monthly, but also required only one representative of each group to participate (Pellegrina & Michele, 2017). Of the five studies identified in the pilot study, all five reported null associations between ‘more regular’ meetings and loan repayment, although the baseline regularity of meetings varies between studies. Of the two studies from the main literature search identified as having comparable predictor, outcome,

and effect measures (Berg et al., 2015; Singh & Padhi, 2017), both showed a null association between more regular meetings and loan repayment.

Sensitivity analyses of our priors indicated that our findings are robust to alternative choices of priors for μ and τ (Table S4). Meta-regressions including various moderators (number of study covariates, group size, borrower age, and borrower sex) failed to show better performance than our original meta-analytic models reported above with no moderators (Table S5). Generalisability estimates using between-study variance yielded mixed results, with 5 of the 11 variables classified as having ‘reasonable generalisability’, 4 as having ‘low generalisability’, and 2 as having ‘very low generalisability’ (Table S6). Finally, there was little indication of publication bias, with no variable’s value of Bai’s D exceeding 0.1 (Table S7), although this is perhaps unsurprising given that none of our meta-analyses demonstrated clear evidence of an effect of the variable on loan repayment.

Discussion

The global popularity of joint liability microfinance loans enabled us to create a cross-cultural dataset on human cooperation in the form of group loan repayment, featuring a range of diverse, non-WEIRD (Henrich et al., 2010) societies. Following Lamba (2014) and Gehrig et al. (2021), we used this dataset to test predictions derived from evolutionary theories of cooperation, given that joint liability microfinance repayment constitutes a cooperative dilemma. We meta-analysed the associations between 11 variables of evolutionary interest and successful loan repayment. To our knowledge, this is the first attempt to quantitatively meta-analyse the microfinance loan repayment literature informed by evolutionary theories of cooperation. We discuss our results grouped by their associated predictions.

Common ancestry and kin selection

We predicted that loan groups composed of a higher proportion of genetic relatives would exhibit higher repayment rates. There was weak evidence against this prediction from our meta-analysis, which provided a posterior distribution indicating a 65% probability that loan groups with a higher proportion of relatives have lower odds of successfully repaying their loan. This result should be interpreted with caution, as the estimate is uncertain, and the analysis included only 2 studies out of the 18 that we identified in the microfinance literature that include data on the relationship between borrower relatedness and repayment.

Our results suggest repayment in loan groups is not driven by kin-directed cooperation, and that there might even be a negative association between borrower relatedness and repayment. Gehrig et al. (2021) similarly found predominantly negative or null associations between the presence of relatives in loan groups and repayment outcomes. Gehrig et al. (2021) proposed three potential explanations for this: the difficulty of enforcement between relatives due to shared genetic interests (and therefore additional indirect costs of punishment), prioritisation of the needs of additional family members outside the loan group over the interests of the group, and populations with a high density of relatives leading to kin competition that overrides cooperation at the level of the loan group.

A related potential explanation is that punishment between relatives is too costly, not because of indirect fitness costs but due to the likelihood of regular future interaction with related group members when living in kin-structured populations, in which many cooperative behaviours are driven by kin-based reciprocity (Power & Ready, 2019). If punishing a related defaulter in one’s loan group risks damaging a valuable family relationship, the trade-offs to punishment may instead favour inaction, especially if the consequences of their default (e.g. requiring additional repayment contributions) are shared among the loan group. Default in the kin-based loan groups may therefore reflect the prioritisation of reciprocity-based relationships with kin over group loan repayment. This explanation, and its focus on future reciprocated benefits rather than genetic fitness interests, aligns with other

studies which have found that across human societies, reciprocity is generally a stronger driver of cooperation than kin selection (Allen-Arave *et al.*, 2008; Jaeggi & Gurven, 2013; Kaplan *et al.*, 1985).

The uncertainty of the estimate for relatedness may also be due to the inconsistent and poorly specified definitions of ‘relatives’ in the microfinance literature. The two studies that we used in our meta-analysis defined relatedness as the ‘percent of group members having a close relative in the group’ (Ahlin & Townsend, 2007, p.F31) and the ‘percentage of group peers who are relatives of the respondent’ (Qinlan & Izumida, 2013, p.334), without specifying exactly who is classified as a relative vs a non-relative or a close vs distant (or non-)relative. Classic research in the evolutionary human sciences has demonstrated that cooperative behaviour differs towards genetically related family members vs non-genetically related family members (Daly & Wilson, 1988) and that cooperation is moderated by the exact degree of genetic relatedness between individuals (Burnstein *et al.*, 1994; Silk, 1980). While some microfinance studies explicitly identify ‘blood relatives’ (Chowdhury, 2016) or record ‘consanguinity’ relationships (Chen *et al.*, 2020), which could be interpreted as implying genetic relatedness, such instances are rare and still rather imprecise. Future microfinance research would benefit from recording the exact degree of genetic relatedness between group members to improve our understanding of how relatedness impacts loan repayment.

Prior interaction and reciprocity

We predicted that loan groups composed of borrowers with experience or knowledge of one another’s past cooperative behaviour would facilitate reciprocal cooperation and therefore lead to higher repayment rates. Our meta-analytic results provide three independent tests of this prediction.

First, there was weak supportive evidence for reciprocity from a meta-analysis indicating a 65% probability that groups with previously acquainted borrowers have higher odds of successfully repaying their loan. However, the posterior distribution for this estimate peaks around a null association, and the estimate is uncertain. The data for this analysis were unusual in containing a near 50/50 split between studies, suggesting a significant negative association between borrower acquaintance and repayment, and those suggesting a significant positive association, resulting in a high level of between-study heterogeneity. The level of analysis of contributing studies appears to partly explain this heterogeneity (Table S4), with acquaintance between borrowers associated with better individual-level repayment performance (OR 2.71, 95% CI 0.59–9.76), but worse group-level repayment performance (OR 0.66, 95% CI 0.26–1.93). This finding is hard to reconcile, given that group repayment is a function of individual borrower repayment. It could suggest that existing social ties within a group benefit individuals at the expense of the group, for example, by acquainted members showing leniency for repayment deadlines so that group performance is rated poorly due to delinquency, but individual borrowers ultimately repay their loans. As with cooperation between relatives, reciprocal cooperation may be harder to sustain between acquaintances if the expected value of a relationship means the costs of punishment or risks of retaliation are high, so that poor repayment discipline is tolerated (Wyduck, 1999). However, CI for both estimates include null associations, and it is also possible that this result is an artefact of sampling variation, or non-random differences between studies using different levels of measurement for loan repayment.

Second, there was neither evidence for nor against reciprocity from a meta-analysis, indicating a 56% probability that groups with longer tenure have lower odds of successfully repaying their loan. The median of the posterior distribution for this estimate lay just below 1, with a long right tail in the distribution. This is most consistent with a null or very small negative association between the length of time a loan group has existed and its repayment performance. A caveat to this analysis is the risk of selection bias in the measurement of group tenure – loan groups that fail early in their existence due to a lack of time for reciprocity-based relationships to form are inherently less likely to be included in studies of loan repayment, thus distorting estimates of the effect of group tenure.

Third, there was some evidence against reciprocity from a meta-analysis indicating a 75% probability that borrowers who live or work in closer geographic proximity have lower odds of successfully repaying their loan, with a mean estimate of 7% lower odds of repayment for each kilometre increase in proximity. This result contradicts theoretical predictions that group members who are better able to monitor one another should display better repayment performance (Besley & Coate, 1995). However, geographic proximity between borrowers may not be an accurate proxy of monitoring propensity, and it may be correlated with other environmental factors that influence repayment (Karlan, 2007). Another possible interpretation is that for individuals to enter a group with geographically distant partners, there must be strong pre-existing levels of trust, leading to higher repayment rates.

Our findings regarding the role of reciprocity in loan groups parallel those of Gehrig et al. (2021), which showed mixed evidence, with 64% of contributing studies showing a negative or non-significant association between prior borrower interaction and loan repayment. A more precise test of the role of reciprocity within loan groups could be achieved with longitudinal data capable of identifying temporal patterns of individual repayment in response to the repayment decisions of fellow group members. If reciprocity between borrowers underlies repayment in group lending, it could be responsible for both higher and lower repayment rates. We would therefore predict individual default in one repayment round to cause default by other group members in the next round, in a tit-for-tat pattern (Axelrod & Hamilton, 1981; Trivers, 1971). In an analysis of group lending data from an Indian lending institution, borrowers were 10–15 percentage points more likely to repay their loans when all other group members repaid (Breza, 2012). However, this effect was asymmetric in that successful peer repayment promoted individual repayment, but peer default was not followed by an increase in individual default.

Partner choice

We predicted that loan groups formed by borrowers themselves, and those which actively manage their membership through screening, would exhibit higher repayment rates. This prediction was unsupported by a meta-analysis indicating a 54% probability that groups which manage their membership have higher odds of successfully repaying their loan. The posterior distribution for this estimate peaks around a null association. Between-study heterogeneity for this analysis was fairly extreme, which may reflect variation in the choice of measurements of self-formation and screening of new borrowers. The estimate provided by Asgedom et al. (2015) is negative but also highly uncertain, and the only statistically significant negative estimate included is from the only experimental study, in which borrowers participated in a lab-in-the-field microfinance game (Cassar et al., 2007). However, the game played in this study had questionable external validity, given that repayment decisions in the game resulted in returns to borrowers that were kept private. It is the influence of these two studies, which draws the meta-analytic estimate down. Still, our results generally reflect the findings of Gehrig et al. (2021), who found a greater number of studies providing statistically significant positive associations than negative associations between partner choice and group loan repayment.

It is surprising that there is not stronger evidence of a positive association for partner choice, given that borrowers have information on the cooperativeness of potential partners. However, no studies measured ongoing partner choice within the lifetime of the loan group, such as whether a group had ever expelled or replaced a borrower for poor repayment performance. We predict that if partner choice was measured in this way, its association with successful loan repayment would be stronger.

Punishment

We predicted that loan groups, which carried out sanctions against defaulters, and groups, which conducted peer or external monitoring, would exhibit higher repayment rates. Our results provide three tests of this prediction.

First, this prediction was weakly supported by a meta-analysis indicating a 90% probability that groups which sanction defaulters have higher odds of successfully repaying their loan, with a median estimate of 89% greater odds of repayment. This estimate is uncertain and displays high levels of between-study heterogeneity, although this is mostly driven by variation among positive estimates from the original studies. While this suggests a positive overall association between punishment and repayment, all studies contributing to our analysis used measures of potential punishment, rather than measuring instances in which punishment for non-repayment of loans actually occurred. This proxy measure is common in studies of human punishment, although the assumption that the threat of punishment functions to convert would-be free-riders into cooperators is debatable (Raihani & Bshary, 2019). Evidence suggests instead that punishment is motivated by inequity aversion, whereby individuals are sensitive to differential outcomes between themselves and their interaction partners (Fehr & Schmidt, 1999; Raihani & McAuliffe, 2012) and use punishment to alter their relative pay-offs. This is supported by an international analysis of microfinance repayment, which concludes that higher levels of inequity aversion among borrowers predict increased default rates (Jordan *et al.*, 2019), as groups refuse to repay the dues of members who are perceived to have profited unfairly through strategic default.

As noted in the Introduction section, little research has examined actual mechanisms of punishment within joint liability loan groups, beyond anecdotal evidence of social shaming, confrontation and confiscation of possessions (Solli *et al.*, 2015). The group-based nature of such punishment behaviour provides support for theories that focus on the sharing of punishment costs (Boyd *et al.*, 2010, 2003). Future research should explore the relative costs and benefits to individual members involved in microfinance punishment scenarios, as punishment can only evolve as a mechanism if there is a correlation between punishment behaviour and the benefits of cooperation which the punisher receives (Gardner & West, 2004).

Second, this prediction was supported by a meta-analysis indicating a 94% probability that groups which conduct peer monitoring have higher odds of successfully repaying their loan. This estimate is still fairly uncertain, although more certain than the estimate for group sanctions. Some of this uncertainty may be due to variation in the variables used as proxy measures of monitoring, including the regularity of visits between members, and knowledge of other group members' business activities. Peer monitoring is one of the main functions of the loan group and the majority of peer monitoring may occur discreetly during dedicated events, such as repayment meetings in which members report their investment returns.

Third, there was some evidence against this prediction from a meta-analysis indicating an 87% probability that groups that undergo external monitoring have lower odds of successfully repaying their loan. This estimate is more certain, although the posterior distribution has a long right tail. While it may seem counterintuitive that externally monitored groups show worse repayment rates, this finding may be due to reverse causality. Most MFIs do not have the resources to conduct monitoring of groups, reserving inspection visits for groups with historical repayment problems (Armendáriz & Morduch, 2010).

Group size, borrower age, and borrower sex

Due to their regular inclusion as covariates in studies of loan repayment, the meta-analyses of group size, borrower age, and borrower sex had the largest sample sizes. Hence, their estimates showed the greatest certainty. Our analysis provides some evidence that larger groups are associated with improved repayment performance, contradicting theoretical predictions that larger groups should make the establishment of reciprocal relationships more difficult (Boyd & Richerson, 1988). However, this association between group size and cooperation is unlikely to be linear, and may only apply within a certain range: group sizes in our dataset ranged from 2 to 18, which may reflect upper limits on borrowers' ability to track cooperative behaviours.

Borrower age provided the most certain estimate of all our analyses, although the magnitude of the association is close to null, likely because the association represents the difference in repayment from only a single year increase in age. There is stronger evidence of a positive association between repayment and being female, which supports the common assumption that women generally show better repayment performance (Armendáriz & Morduch, 2010), possibly because they tend to be more risk-averse in their loan investment choices or more responsive to the threat of social shaming. The 63% female makeup of our dataset reflects a historical focus on women in microfinance, particularly in joint liability lending (D'Espallier et al., 2013) and is consistent with a global analysis of 350 lending institutions, which found 73% of borrowers were female (D'Espallier et al., 2011).

Limitations

Several limitations necessitate caution in the interpretation of our results, especially for those considering their practical application in a novel lending context. Restricting our sample to comparable effect estimates – a prerequisite for any meta-analysis – reduced our meta-analytic sample sizes, so that for each variable we were only able to analyse between 11% and 57% of estimates available in the literature. This was further exaggerated in our subgroup and sensitivity analyses. All of our meta-analyses constitute univariate analyses as we were unable to conduct multivariate meta-regressions capable of controlling for known confounds (Table S5), largely due to the lack of shared covariate sets between studies.

Our study findings may also be affected by endogeneity in loan group formation, whereby economic or environmental conditions cause groups to form in a certain way (Berhane et al., 2009). For example, borrowers may be more likely to form groups with family members during an economic downturn. Assuming such an economic downturn would also increase default rates due to poor business performance, this could result in a correlation between borrower relatedness and default, even though, in a counterfactual scenario, borrowers may have had even higher default rates had they not formed groups with family members. Alternatively, some variables such as group size may not be exogenous to the maintenance of cooperation. Large groups may be large precisely because some other mechanism is already in place that facilitates cooperation, causing more people to join that group. This endogeneity may explain the theoretically unexpected finding that larger groups are associated with higher repayment. Additionally, there is a risk of site selection bias among studies of loan repayment, if lending institutions that collaborate with researchers differ in some non-random way from institutions that are not engaged in research, such as having better repayment performance (Allcott, 2015; Reid et al., 2018).

Our eligible studies display considerable variation in study settings, including variation in features of joint liability loan contracts that could influence repayment decisions. For example, interest rates vary within and between lending institutions, and higher or lower rates could influence the costs and benefits of loan repayment for borrowers. Loan features such as interest rates and loan size could also correlate with higher-order factors associated with loan repayment outcomes, such as the capacity or experience of the lending institution, or adverse selection of riskier borrowers (Aleem, 1990; Bottomley, 1975). We could not control for such factors in our analyses.

There is variation in the measurement of predictor and outcome variables, even among comparable studies. When combining predictor variables, we aimed to strike a balance between identifying functionally similar variables and maximising the sample sizes available for our analyses, but due to differences in expert judgement there are likely to be reasonable objections to some of our categorisation decisions. To improve the generalisability of mechanisms promoting cooperation and loan repayment, we encourage broader collaboration between microfinance researchers, particularly joint efforts to collect data on the same variables for loan groups borrowing from different lending institutions, and to test the same theoretical models of repayment in different study contexts (Vivalt, 2020).

Loan repayment studies which measure only whether loans were ultimately repaid at the group level may be overlooking the most interesting data on cooperation between borrowers: the proportion of successfully repaid group loans that involved any instance of individual default. Groups that face individual default but which successfully overcome this through fully repaying their loan have arguably engaged in a more costly act of cooperation than groups in which all borrowers repay their dues. Conversely, studies which measure only individual repayment performance cannot assess whether the cooperative dilemma of group lending was solved. While some studies do report individual- and group-level delinquency separately (Bassem, 2008), no studies report the combined results. This mirrors a measurement problem among lending institutions, which tend to measure loan repayment only at the institutional level (Rosenberg, 1999), reflecting their lack of concern regarding *how* group loans are repaid – only *that* they are repaid.

Finally, almost all of our estimates are drawn from cross-sectional studies, limiting our ability to infer causal relationships between our predictors of interest and loan repayment. While our meta-analytic estimates provide indirect evidence of causal relationships, and have value in predicting loan repayment outcomes, they should not be interpreted as causal effects. Our understanding of the causal factors that promote cooperation and repayment in group loans would benefit from more studies using randomisation-based designs in natural settings, such as field-based RCTs, enabling manipulation of variables of interest between different treatment groups. There are notable examples of high-quality RCTs in microfinance research (Field *et al.*, 2012; Giné & Karlan, 2014), but these tend to focus on the effects of manipulating features of the loan contract, such as liability structure, rather than manipulating features of joint liability groups likely to facilitate or inhibit cooperation amongst their members.

Conclusion

Our meta-analyses do not provide conclusive evidence regarding the role of evolutionary mechanisms of cooperation supporting loan repayment in joint liability groups. We find little support for kin selection, reciprocity, and partner choice. The two variables showing the strongest evidence of a positive association with loan repayment are both related to punishment. The predominance of uncertain associations in our results may be due to statistical variation between studies due to imprecise estimates or may reflect true variation in the role of evolutionary mechanisms across different cultural and ecological contexts.

Supplementary material. The supplementary material for this article can be found at <https://doi.org/10.1017/ehs.2026.10047>.

Acknowledgements. We thank Rachael Meager, Lucia Dalla Pellegrina, and Bruce Wydick for feedback on the study protocol; Wolfgang Viechtbauer, Mark Kelson, Robert Grant, and Ariel Aloe for meta-analysis statistical advice; Ailsa Poll for literature search advice and library support; Luke Arundel for second-coding study effect estimates; and Charles Efferson, Scott Claessens, and one anonymous reviewer for valuable comments on the manuscript.

Author contributions. DF, SL, EP, and AM designed the study and co-wrote the pre-registration. DF conducted the pilot study, data collection, and analyses, with additional analyses by AM. DF wrote the first draft of the paper, with input and editing by SL, EP, and AM.

Financial support. This work was funded by a Royal Society fellowship (UF150036) and grant (RG160537) to SL.

Competing interests. The authors declare no conflicts of interest.

Research transparency and reproducibility interest. All data and code are available at <https://github.com/dugaldfoster/MicrofinanceMetaAnalysis> and <https://osf.io/wsdjn/>.

References

- Ågren, J. A., Davies, N. G., & Foster, K. R. (2019). Enforcement is central to the evolution of cooperation. *Nature Ecology and Evolution*, 3(7), 1018–1029. <https://doi.org/10.1038/s41559-019-0907-1>
- Ahlin, C., & Townsend, R. M. (2007). Using repayment data to test across models of joint liability lending. *The Economic Journal*, 117(517), F11–F51. <https://doi.org/10.1111/j.1468-0297.2007.02014.x>
- Aktipis, C. A. (2004). Know when to walk away: Contingent movement and the evolution of cooperation. *Journal of Theoretical Biology*, 231(2), 249–260. <https://doi.org/10.1016/j.jtbi.2004.06.020>
- Aleem, I. (1990). Imperfect information, screening, and the costs of informal lending: A study of a rural credit market in Pakistan. *The World Bank Economic Review*, 4(3), 329–349. <https://doi.org/10/fs58zt>
- Allcott, H. (2015). Site selection bias in program evaluation. *The Quarterly Journal of Economics*, 130(3), 1117–1165. <https://doi.org/10.1093/qje/qjv015>
- Allen-Arave, W., Gurven, M., & Hill, K. (2008). Reciprocal altruism, rather than kin selection, maintains nepotistic food transfers on an Ache reservation. *Evolution and Human Behavior*, 29(5), 305–318. <https://doi.org/10.1016/j.evolhumbehav.2008.03.002>
- Apicella, C. L., & Silk, J. B. (2019). The evolution of human cooperation. *Current Biology*, 29(11), R447–R450. <https://doi.org/10.1016/j.cub.2019.03.036>
- Armendáriz, B. (1999). On the design of a credit agreement with peer monitoring. *Journal of Development Economics*, 60(1), 79–104. [https://doi.org/10.1016/S0304-3878\(99\)00037-1](https://doi.org/10.1016/S0304-3878(99)00037-1)
- Armendáriz, B., & Morduch, J. (2010). *The Economics of Microfinance*. MIT Press.
- Asgedom, A. K., Desta, T. S., & Bahita, G. H. (2015). Factors affecting group loan repayment performance: A case of debit credit and saving institution (DECSI), Mekelle, Ethiopia. *Journal of Poverty, Investment and Development*, 10(1), 1–22. (<https://www.iiste.org/Journals/index.php/JPID/article/view/20849>)
- Attanasio, O., Augsburg, B., Haas, R. D., Fitzsimons, E., & Harmgart, H. (2015). The impacts of microfinance: Evidence from joint-liability lending in Mongolia. *American Economic Journal: Applied Economics*, 7(1), 90–122. <https://doi.org/10.1257/app.20130489>
- Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science*, 211(4489), 1390–1396. <https://doi.org/10.1126/science.7466396>
- Bai, R., Lin, L., Boland, M. R., & Chen, Y. (2020). A robust Bayesian Copas selection model for quantifying and correcting publication bias. arXiv preprint arXiv:2005.02930. <https://doi.org/10.48550/arXiv.2005.02930>
- Banerjee, A., Karlan, D., & Zinman, J. (2015). Six randomized evaluations of microcredit. *American Economic Journal: Applied Economics*, 7(1), 1–21. <https://doi.org/10.1257/app.20140287>
- Bassem, B. S. (2008). Determinants of successful group loan repayment: An application to Tunisia. *Journal of Sustainable Development in Africa*, 10(2), 766–800. (https://jsd-africa.com/Jsda/v10n2_Summer2008/PDF/ARC_Bassem%20and%20Borhen-%20Determinants.pdf)
- Bastelaer, T. V., & Leathers, H. (2006). Trust in lending: Social capital and joint liability seed loans in Southern Zambia. *World Development* 34(10), 1788–1807. <https://doi.org/10.1016/j.worlddev.2006.02.007>
- Berg, M., van den Lensink, R., & Servin, R. (2015). Loan officers' gender and microfinance repayment rates. *The Journal of Development Studies* 51(9), 1241–1254. <https://doi.org/10/ggdnsz>
- Berhane, g, Gardebroeck, c, & Moll, H. A. (2009). Risk-matching behavior in microcredit group formation: evidence from northern Ethiopia. *Agricultural Economics*, 40(4), 409–419. <https://doi.org/10.1111/j.1574-0862.2009.00388.x>
- Besley, T., & Coate, S. (1995). Group lending, repayment incentives and social collateral. *Journal of Development Economics*, 46(1), 1–18. [https://doi.org/10.1016/0304-3878\(94\)00045-E](https://doi.org/10.1016/0304-3878(94)00045-E)
- Bottomley, A. (1975). Interest rate determination in underdeveloped rural areas. *American Journal of Agricultural Economics*, 57(2), 279–291. <https://doi.org/10/ckmtfj>
- Bowles, S., & Gintis, H. (2004). The evolution of strong reciprocity: Cooperation in heterogeneous populations. *Theoretical Population Biology*, 65(1), 17–28. <https://doi.org/10.1016/j.tpb.2003.07.001>
- Boyd, R., Gintis, H., & Bowles, S. (2010). Coordinated punishment of defectors sustains cooperation and can proliferate when rare. *Science*, 328(5978), 617–620. <https://doi.org/10.1126/science.1183665>
- Boyd, R., Gintis, H., Bowles, S., & Richerson, P. J. (2003). The evolution of altruistic punishment. *Proceedings of the National Academy of Sciences*, 100(6), 3531–3535. <https://doi.org/10.1073/pnas.0630443100>
- Boyd, R., & Richerson, P. J. (1988). The evolution of reciprocity in sizable groups. *Journal of Theoretical Biology*, 132(3), 337–356. [https://doi.org/10.1016/S0022-5193\(88\)80219-4](https://doi.org/10.1016/S0022-5193(88)80219-4)
- Boyd, R., & Richerson, P. J. (1992). Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethology and Sociobiology*, 13(3), 171–195. [https://doi.org/10.1016/0162-3095\(92\)90032-Y](https://doi.org/10.1016/0162-3095(92)90032-Y)
- Bramer, W. M., Rethlefsen, M. L., Kleijnen, J., & Franco, O. H. (2017). Optimal database combinations for literature searches in systematic reviews: A prospective exploratory study. *Systematic Reviews*, 6(1), 245. <https://doi.org/10/ggfcv5>
- Breza, E. (2012). Peer effects and loan repayment: Evidence from the Krishna default crisis. *Job Market Paper MIT*, 1. (https://business.columbia.edu/sites/default/files-efs/pubfiles/6030/Breza_PeerEffects.pdf)

- Brockwell, S. E., & Gordon, I. R. (2001). A comparison of statistical methods for meta-analysis. *Statistics in Medicine*, 20(6), 825–840. <https://doi.org/10/bcp95p>
- Bürkner P. (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. *J. Stat. Soft.*, 80(1). <https://doi.org/10.18637/jss.v080.i01>
- Burnstein, E., Crandall, C., & Kitayama, S. (1994). Some neo-Darwinian decision rules for altruism: Weighing cues for inclusive fitness as a function of the biological importance of the decision. *Journal of Personality and Social Psychology*, 67(5), 773. <https://doi.org/10.1037/0022-3514.67.5.773>
- Carpena, F., Cole, S., Shapiro, J., & Zia, B. (2013). Liability structure in small-scale finance. *The World Bank Economic Review*, 27(3), 437–469. <https://doi.org/10.1093/wber/lhs031>
- Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., & Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, 76(1). <https://doi.org/10/b2pm>
- Carpenter, J., & Williams, T. (2014). Peer monitoring and microcredit: Field experimental evidence from Paraguay. *Oxford Development Studies*, 42(1), 111–135. <https://doi.org/10.1080/13600818.2014.887061>
- Cassar, A., Crowley, L., & Wydick, B. (2007). The effect of social capital on group loan repayment: Evidence from field experiments. *The Economic Journal* 117(517), F85–F106. <https://doi.org/10.1111/j.1468-0297.2007.02016.x>
- Chen, Y., Deng, K., & Xu, Z. (2020). The social capital in microfinance issued by commercial banks: Evidence from China. *Applied Economics Letters* 27(4), 349–352. <https://doi.org/10.1080/13504851.2019.1616054>
- Chowdhury, H. R. (2016). Joint-liability in microcredit: Evidence from Bangladesh. *Atlantic Economic Journal*, 44(1), 105–129. doi:10.1007/s11293-016-9490-x
- Conning, J., & Udry, C. (2007). Rural financial markets in developing countries. In *Handbook of Agricultural Economics* (Vol. 3, pp. 2857–2908). Elsevier. <https://ideas.repec.org/h/eee/hagchp/5-56.html>
- Copas, J. (1999). What works?: Selectivity models and meta-analysis. *Journal of the Royal Statistical Society Series B* 62, 95–109. <https://doi.org/10/fb4m24>
- Copas, J. B., & Shi, J. Q. (2001). A sensitivity analysis for publication bias in systematic reviews. *Statistical Methods in Medical Research* 10, 251–265. <https://doi.org/10/gf495n>
- Cull, R., Ç-kunt, A. D., & Morduch, J. (2007). Financial performance and outreach: A global analysis of leading microbanks. *The Economic Journal* 117(517), F107–F133. <https://doi.org/10.1111/j.1468-0297.2007.02017.x>
- Despallier, B., Guerin, I., & Mersland, R. (2013). Focus on women in microfinance institutions. *The Journal of Development* 49(5), 589–608. <https://doi.org/10/ggwsnv>
- Despallier, B., Guérin, I., & Mersland, R. (2011). Women and repayment in microfinance: A global analysis. *World Development* 39(5), 758–772. <https://doi.org/10.1016/j.worlddev.2010.10.008>
- Daly, M., & Wilson, M. (1988). Evolutionary social psychology and family homicide. *Science*, 242(4878), 519–524. <https://doi.org/10.1126/science.3175672>
- Dawes, R. M. (1980). Social dilemmas. *Annual Review of Psychology*, 31(1), 169–193. <https://doi.org/10.1146/annurev.ps.31.020180.001125>
- Ertan, A., Page, T., & Putterman, L. (2009). Who to punish? Individual decisions and majority rule in mitigating the free rider problem. *European Economic Review*, 53(5), 495–511. <https://doi.org/10.1016/j.euroecorev.2008.09.007>
- Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature*, 415(6868), 137–140. <https://doi.org/10.1038/415137a>
- Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics*, 114(3), 817–868. <https://doi.org/10.1162/003355399556151>
- Feigenberg, B., Field, E., & Pande, R. (2013). The economic returns to social interaction: Experimental evidence from microfinance. *Review of Economic Studies*, 80(4), 1459–1483. <https://doi.org/10.1093/restud/rdt016>
- Field, E., Pande, R., Papp, J., & Park, Y. J. (2012). Repayment flexibility can reduce financial stress: A randomized control trial with microfinance clients in India. *PloS One*, 7(9). <https://doi.org/10/ggjpgvq>
- Fletcher, J. A., & Michael Doebeli, M. (2009). A simple and general explanation for the evolution of altruism. *Proceedings of the Royal Society B* 276(1654), 13–19. doi:10.1098/rspb.2008.0829
- Frank, S. A. (1995). Mutual policing and repression of competition in the evolution of cooperative groups. *Nature*, 377(6549), 520. <https://doi.org/10.1038/377520a0>
- Frank, S. A. (2003). Repression of competition and the evolution of cooperation. *Evolution*, 57(4), 693–705. <https://doi.org/10.1111/j.0014-3820.2003.tb00283.x>
- Gardner, A., & West, S. A. (2004). Spite and the scale of competition. *Journal of Evolutionary Biology*, 17(6), 1195–1203. <https://doi.org/10/bq4cb9>
- Gehrig, S., Mesoudi, A., & Lamba, S. (2021). Banking on cooperation: An evolutionary analysis of microfinance loan repayment behaviour. *Evolutionary Human Sciences* 3,e2. <https://doi.org/10.1017/ehs.2020.64>
- Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian Analysis*, 1(3), 515–534. <https://doi.org/10/bh3qjf>
- Gelman, A. (2018). The failure of null hypothesis significance testing when studying incremental changes, and what to do about it. *Personality and Social Psychology Bulletin*, 44(1), 16–23. <https://doi.org/10.1177/0146167217729162>

- Gelman, A. (2019). *Prior Choice Recommendations* [Stan]. Stan. <https://github.com/stan-dev/stan> (Original work published 2013)
- Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian Data Analysis* (3rd ed). CRC Press.
- Ghatak, M., & Guinnane, T. (1999). The economics of lending with joint liability. *Journal of Development Economics* 60(1), 195–228. [https://doi.org/10.1016/S0304-3878\(99\)00041-3](https://doi.org/10.1016/S0304-3878(99)00041-3)
- Giné, X., & Karlan, D. (2014). Group versus individual liability: Short and long term evidence from Philippine microcredit lending groups. *Journal of Development Economics* 107, 65–83. <https://doi.org/10.1016/j.jdeveco.2013.11.003>
- Grafen, A. (2007). Detecting kin selection at work using inclusive fitness. *Proceedings of the Royal Society B*, 274(1610), 713–719. <https://doi.org/10.1098/rspb.2006.0140>
- Gürerk, Ö., Irlenbusch, B., & Rockenbach, B. (2006). The competitive advantage of sanctioning institutions. *Science*, 312(5770), 108–111. <https://doi.org/10.1126/science.1123633>
- Gurevitch, J., Koricheva, J., Nakagawa, S., & Stewart, G. (2018). Meta-analysis and the science of research synthesis. *Nature*, 555(7695), 175–182. <https://doi.org/10.1038/nature25753>
- Haddaway, N. R., Collins, A. M., Coughlin, D., & Kirk, S. (2015). The role of Google Scholar in evidence reviews and its applicability to grey literature searching. *PLOS ONE*, 10(9), e0138237. <https://doi.org/10/gcpcfhw>
- Haldar, A., & Stiglitz, J. (2016). Group lending, joint liability, and social capital: Insights from the Indian microfinance crisis. *Politics & Society* 44(4), 459–497. <https://doi.org/10/f9b23m>
- Hamilton, W. D. (1964a). The genetical evolution of social behaviour. I. *Journal of Theoretical Biology*, 7(1), 1–16. [https://doi.org/10.1016/0022-5193\(64\)90038-4](https://doi.org/10.1016/0022-5193(64)90038-4)
- Hamilton, W. D. (1964b). The genetical evolution of social behaviour. II. *Journal of Theoretical Biology*, 7(1), 17–52. [https://doi.org/10.1016/0022-5193\(64\)90039-6](https://doi.org/10.1016/0022-5193(64)90039-6)
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Most people are not WEIRD. *Nature*, 466, 29. <https://doi.org/10.1038/466029a>
- Herrmann, B., Thöni, C., & Gächter, S. (2008). Antisocial punishment across societies. *Science*, 319(5868), 1362–1367. <https://doi.org/10.1126/science.1153808>
- Higgins, J. P. T., & Thomas, J. (2019). *Cochrane Handbook for Systematic Reviews of Interventions*. John Wiley & Sons.
- Hoffman, M. D., & Gelman, A. (2014). The no-u-turn sampler: adaptively setting path lengths in hamiltonian monte carlo. *Journal of Machine Learning Research*, 15(47), 1593–1623. (<https://dl.acm.org/doi/abs/10.5555/2627435.2638586>)
- Hruschka, D. J., & Henrich, J. (2006). Friendship, cliquishness, and the emergence of cooperation. *Journal of Theoretical Biology*, 239(1), 1–15. <https://doi.org/10.1016/j.jtbi.2005.07.006>
- Isaac, R. M., & Walker, J. M. (1988). Group size effects in public goods provision. *The Quarterly Journal of Economics*, 103(1), 179–199. JSTOR. <https://doi.org/10/fjbnbc>
- Jaeggi, A. V., & Gurven, M. (2013). Natural cooperators: Food sharing in humans and other primates. *Evolutionary Anthropology*, 22(4), 186–195. <https://doi.org/10.1002/evan.21364>
- Janssens, A. C. J. W., Gwinn, M., Brockman, J. E., Powell, K., & Goodman, M. (2020). Novel citation-based search method for scientific literature. *BMC Medical Research Methodology*, 20(1), 25. <https://doi.org/10/ggkz2z>
- Jordan, M., Dickens, W., Hauser, O., & Rand, D.G. (2019). The role of inequity aversion in microloan defaults. *Behavioural Public Policy*, 6(2), 303–324. <https://doi.org/10.1017/bpp.2019.29>
- Kaplan, H., Hill, K., Cadelina, R. V., Hayden, B., Hyndman, D. C., Preston, R. J., Smith, E. A., Stuart, D. E., & Yesner, D. R. (1985). Food sharing among ache foragers. *Current Anthropology*, 26(2), 223–246. <https://doi.org/10/b8csgf>
- Karlan, D. S. (2007). Social connections and group banking. *The Economic Journal* 117(517), F52–F84. <https://doi.org/10/df4mzh>
- Kerr N L. (1998). HARKing: Hypothesizing After the Results are Known. *Pers Soc Psychol Rev*, 2(3), 196–217. https://doi.org/10.1207/s15327957pspr0203_4
- Lamba, S. (2014). A comparison of the economic literature on microfinance and the evolutionary literature on cooperation. In M. A. Gibson & D. W. Lawson (Eds.), *Applied Evolutionary Anthropology* (pp. 39–57). Springer New York. https://doi.org/10.1007/978-1-4939-0280-4_3
- Lehmann, L., & Keller, L. (2006). The evolution of cooperation and altruism – A general framework and a classification of models. *Journal of Evolutionary Biology*, 19(5), 1365–1376. <https://doi.org/10.1111/j.1420-9101.2006.01119.x>
- Mahmud, M. (2019). Repaying microcredit loans: A natural experiment on liability structure. *The Journal of Development Studies*, 1–16. <https://doi.org/10/ggkphd>
- Maynard Smith, J. (1964). Group selection and kin selection. *Nature*, 201(4924), 1145. <https://doi.org/10.1038/2011145a0>
- McElreath, R. (2016). *Statistical rethinking: A Bayesian course with examples in R and Stan*. CRC Press.
- McNamara, J. M., Barta, Z., Fromhage, L., & Houston, A. I. (2008). The coevolution of choosiness and cooperation. *Nature*, 451(7175), 189–192. <https://doi.org/10.1038/nature06455>
- Meager, R. (2019). Understanding the average impact of microcredit expansions: A Bayesian hierarchical analysis of seven randomized experiments. *American Economic Journal: Applied Economics*, 11(1), 57–91. <https://doi.org/10/gfxm4v>

- Microfinance Barometer. (2019). Microfinance Barometer 2019. In *Convergences World Forum* (Vol. 14). https://www.convergences.org/wp-content/uploads/2019/09/Microfinance-Barometer-2019_web-1.pdf
- Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P. & Stewart, L. A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews*, 4(1), 1.
- Morduch, J. (1999). The microfinance promise. *Journal of Economic Literature*, 37(4), 1569–1614. <https://doi.org/10/cp44tf>
- Moreau, D., & Gamble, B. (2022). Conducting a meta-analysis in the age of open science. *Psychological Methods*, 27, 426–432. <https://doi.org/10.1037/met0000351>
- Muka, T., Glisic, M., Milic, J., Verhoog, S., Bohlius, J., Bramer, W., Chowdhury, R., & Franco, O. H. (2020). A 24-step guide on how to design, conduct, and successfully publish a systematic review and meta-analysis in medical research. *European Journal of Epidemiology*, 35(1), 49–60. <https://doi.org/10/ggts4n>
- Noë, R., & Hammerstein, P. (1994). Biological markets: Supply and demand determine the effect of partner choice in cooperation, mutualism and mating. *Behavioral Ecology and Sociobiology*, 35(1), 1–11. <https://doi.org/10.1007/BF00167053>
- Nowak, M. A. (2006). Five rules for the evolution of cooperation. *Science*, 314(5805), 1560–1563. <https://doi.org/10.1126/science.1133755>
- Nowak, M. A., & Sigmund, K. (1998). Evolution of indirect reciprocity by image scoring. *Nature*, 393(6685), 573. <https://doi.org/10.1038/31225>
- Olson, M. (1971). *The logic of collective action*. Harvard University Press.
- Ostrom, E., Walker, J., & Gardner, R. (1992). Covenants with and without a sword: Self-governance is possible. *The American Political Science Review*, 86(2), 404–417. JSTOR. <https://doi.org/10.2307/1964229>
- Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan—A web and mobile app for systematic reviews. *Systematic Reviews*, 5(1), 210. <https://doi.org/10/gfkdzd>
- Pellegrina, L. D., & Michele, A. D., Di Maio, G., Landoni, P., & Parravicini, S. (2017). Group meeting frequency and borrowers' repayment performance in microfinance: Evidence from a quasi-natural experiment in South Africa. University of Milan. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3080201
- Polson, N. G., & Scott, J. G. (2012). On the half-cauchy prior for a global scale parameter. *Bayesian Analysis*, 7(4), 887–902. <https://doi.org/10/f4gw2s>
- Power, E. A., & Ready, E. (2019). Cooperation beyond consanguinity: Post-marital residence, delineations of kin, and social support among South Indian Tamils. *Philosophical Transactions of the Royal Society B*. <https://doi.org/10.1098/rstb.2018.0070>
- Qinlan, Z., & Izumida, Y. (2013). Determinants of repayment performance of group lending in China: Evidence from rural credit cooperatives' program in Guizhou province. *China Agricultural Economic Review*, 5(3), 328–341. <https://doi.org/10.1108/CAER-08-2012-0083>
- Quintana, D. S. (2015). From pre-registration to publication: A non-technical primer for conducting a meta-analysis to synthesize correlational data. *Frontiers in Psychology*, 6, 1549. <https://doi.org/10.3389/fpsyg.2015.01549>
- R Core Team. (2020). *R: A Language and Environment for Statistical Computing* (4.0.3). R Foundation for Statistical Computing. <https://www.R-project.org/>
- Raihani, N. J., & Bshary, R. (2019). *Punishment: One tool, many uses*. *Evolutionary Human Sciences*, 1. <https://doi.org/10.1017/ehs.2019.12>
- Raihani, N. J., & McAuliffe, K. (2012). Human punishment is motivated by inequity aversion, not a desire for reciprocity. *Biology Letters*, 8(5), 802–804. <https://doi.org/10.1098/rsbl.2012.0470>
- Reid, J. L., Fagan, M. E., & Zahawi, R. A. (2018). Positive site selection bias in meta-analyses comparing natural regeneration to active forest restoration. *Science Advances*, 4(5), eaas9143. <https://doi.org/10.1126/sciadv.aas9143>
- Roodman, D. (2012). *Due diligence: An impertinent inquiry into microfinance*. Center for Global Development. <https://muse.jhu.edu/pub/460/monograph/book/86621>
- Rosenberg, R. (1999). *Measuring microcredit delinquency: Ratios can be harmful to your health* (No. 33460, p. 1). The World Bank.
- Salanti, G., Caldwell, D., Chaimani, A., & Higgins, J. (2019). Network Meta-analysis. In A. Levy, S. Goring, C. Gatsonis, B. Sobolev, E. van Ginneken & R. Busse (Eds.), *Health services evaluation* (pp. 577–615). Springer US. https://doi.org/10.1007/978-1-4939-8715-3_36
- Silk, J. B. (1980). Adoption and kinship in Oceania. *American Anthropologist*, 82(4), 799–820. <https://doi.org/10.1525/aa.1980.82.4.02a00050>
- Singh, V., & Padhi, P. (2017). Dynamic incentives and microfinance borrowers: A comparative study of self-help groups and joint liability groups. *Journal of Land and Rural Studies* 5(1), 67–92. <https://doi.org/10/gk4dj9>
- Smaldino, P. E., Schank, J. C., & McElreath, R. (2013). Increased costs of cooperation help cooperators in the Long Run. *The American Naturalist*, 181(4), 451–463. <https://doi.org/10.1086/669615>
- Solli, J., Galindo, L., Rizzi, A., Rhyne, E., & van de Walle, N. (2015). *What happens to microfinance clients who default? The smart campaign* January. https://www.centerforfinancialinclusion.org/wp-content/uploads/2024/02/what_happens_to_microfinance_clients_who_default_eng.pdf
- Spiegelhalter, D. J., Abrams, K. R., & Myles, J. P. (2004). *Bayesian approaches to clinical trials and health care evaluation*. Wiley.

- Stiglitz, J. E. (1990). Peer monitoring and credit markets. *The World Bank Economic Review*, 4(3), 351–366. <https://doi.org/10.1093/wber/4.3.351>
- Suzuki, S., & Akiyama, E. (2005). Reputation and the evolution of cooperation in sizable groups. *Proceedings of the Royal Society B: Biological Sciences*, 272(1570), 1373–1377. <https://doi.org/10/b8vk29>
- RStudio Team (2020) *RStudio: Integrated Development Environment for R* (1.1.447). Rstudio, Inc. <http://www.rstudio.com/>
- Trivers, R. L. (1971). The evolution of reciprocal altruism. *The Quarterly Review of Biology*, 46(1), 35–57. <https://doi.org/10.1086/406755>
- Van Tassel, E. (1999). Group lending under asymmetric information. *Journal of Development Economics*, 60(1), 3–25. <https://doi.org/10/fp3jm3>
- Varian, H. R. (1990). Monitoring agents with other agents. *Journal of Institutional and Theoretical Economics (Jite)/zeitschrift Für Die Gesamte Staatswissenschaft*, 146(1), 153–174. JSTOR. (<https://www.jstor.org/stable/40751313>)
- Vivald, E. (2020). How much can we generalize from impact evaluations? *Journal of the European Economic Association*. <https://doi.org/10/ghg338>
- West, S. A., Griffin, A. S., & Gardner, A. (2007a). Social semantics: Altruism, cooperation, mutualism, strong reciprocity and group selection. *Journal of Evolutionary Biology*, 20(2), 415–432. <https://doi.org/10/b9rmmm>
- West, S. A., Griffin, A. S., & Gardner, A. (2007b). Evolutionary explanations for cooperation. *Current Biology*, 17(16), R661–R672. <https://doi.org/10.1016/j.cub.2007.06.004>
- Wiessner, P. (2005). Norm enforcement among the Ju/'hoansi Bushmen: A case of strong reciprocity? *Human Nature*, 16(2), 115–145. <https://doi.org/10/fhszgm>
- Williams, D. R., Rast, P., & Bürkner, P.-C. (2018). Bayesian meta-analysis with weakly informative prior distributions. *PsyArXiv*. <https://doi.org/10/ggj3rq>
- Woolcock, M. J. V. (1999). Learning from failures in microfinance. *American Journal of Economics and Sociology*. <https://doi.org/10/ft8gbm>
- Wydick B. (1999). Can Social Cohesion be Harnessed to Repair Market Failures? Evidence from Group Lending in Guatemala. *The Economic Journal*, 109(457), 463–475. <https://doi.org/10.1111/1468-0297.00457>
- Yamagishi, T. (1986). The provision of a sanctioning system as a public good. *Journal of Personality and Social Psychology*, 51, 110–116. <https://doi.org/10.1037/0022-3514.51.1.110>