



International Journal of Criminology and Sociological Studies

Journal homepage: <https://trpub.org/journal.php?id=GwyuDBwGBw>

TRPub

IJCSS

International Journal of
Criminology and
Sociological Studies

Vol. 1 Issue 1 2026

Methodology Article

Open Access

Using Respondent-Driven Sampling to Study Hidden Populations of Informal Money Lenders

Lakshmi Al-Rashid^{1*}, Linh Forde², Andre Adeyemi³

¹Department of Social Sciences, Levantine Graduate Institute of Social Sciences, Portugal

²Department of Social Sciences, Mekong Graduate Institute of Social Sciences, Belgium

³Department of Social Sciences, Bharathapuzha Academy of Social Sciences, New Zealand

Article Info

Volume 1, Issue 1, 2026

[doi: 10.62715/ijcss.2026.v1.i1.pp20-25](https://doi.org/10.62715/ijcss.2026.v1.i1.pp20-25)

Received: 16 September 2025

Accepted: 04 November 2025

Published: 10 January 2026

Abstract

Informal money lenders supply a large share of household credit in many low-income economies and operate illegally in many richer ones, yet almost everything known about them comes from their borrowers or from a few ethnographies. They cannot be sampled conventionally because no list of them exists and many have reason to avoid notice. Respondent-driven sampling, a form of chain-referral sampling in which respondents recruit a limited number of peers and report the size of their own networks, was devised for populations of this kind. This article proposes how the method could be adapted to lenders. It reviews the assumptions on which the estimators rest and considers where the social organisation of lending is likely to violate them: lenders compete, many work within hierarchies of financiers and collectors, and the population is divided into types with few ties between them. A design is then set out covering formative research, the definition of eligibility, the choice of seeds, coupons and incentives, the wording of the network-size question, interviewing, estimation and diagnostics. The procedure is a proposal and has not been validated in the field. The article ends by stating what a pilot would have to show before estimates from such a study could be trusted.

Keywords: *Respondent-driven sampling, Hidden populations, Informal credit, Money lenders, Chain-referral sampling, Survey methodology*

© 2026 Lakshmi Al-Rashid et al. This is an open access article under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

1. Introduction

Borrowing from a money lender is an ordinary part of economic life for poor households. Financial diaries kept by families in Bangladesh, India and South Africa record frequent small loans from neighbours, shopkeepers and professional lenders alongside whatever formal credit is available (Collins et al., 2009). Economists have studied the terms of such loans closely. Classic work on rural credit explained high informal interest rates by the costs of screening borrowers and enforcing repayment where information is poor (Hoff and Stiglitz, 1990), and Aleem (1990) measured those costs directly for a set of lenders in one market in Pakistan. In high-income countries the unlicensed lender, the loan shark of popular usage, is a recurring concern of consumer regulators, and in poor urban districts lending shades

* Corresponding author: Lakshmi Al-Rashid, Department of Social Sciences, Levantine Graduate Institute of Social Sciences, Portugal. E-mail: lakshmi.al-rashid@levantine.edu

© 2026 Lakshmi Al-Rashid et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

into the wider off-the-books economy described by [Venkatesh \(2006\)](#).

Studies that sample lenders themselves are rare. Lenders appear in the literature as counterparts in borrowers' accounts, as case material, or as small convenience samples reached through a trusted intermediary. As a result basic descriptive questions remain open in most settings: what kinds of people lend, how many borrowers each serves, where their capital comes from, and how they are linked to one another and to formal finance.

This article considers whether respondent-driven sampling (RDS) could help to answer them. RDS was developed for populations that have no sampling frame and whose members value privacy, and it has become a standard tool in public health surveillance. The article proposes an adaptation to informal lenders and examines its weak points. It does not report a survey. The design is offered for criticism and piloting, and it has not been validated.

2. Background: Respondent-Driven Sampling in Outline

Chain-referral methods are old. Snowball sampling reaches hidden groups by asking each respondent to name others, and [Biernacki and Waldorf \(1981\)](#) set out its practical problems, chief among them that the sample reflects the starting points and the sociability of those recruited. [Heckathorn \(1997\)](#) introduced RDS to deal with these biases. A small number of initial respondents, called seeds, are each given a fixed number of uniquely numbered coupons to pass to peers. Those who come forward are interviewed, rewarded both for taking part and for each peer they later bring in, and given coupons in turn. Because recruitment is done by peers, respondents never have to name anyone to the researcher. Because coupons are limited, no one person can dominate the sample, and chains are pushed to grow long.

Two further elements turn the sample into a basis for inference. The coupon numbers record who recruited whom, and each respondent is asked how many members of the population he or she knows, a quantity called the degree. Under a set of assumptions, the composition of the sample converges to an equilibrium that does not depend on the seeds, and each respondent's probability of inclusion is proportional to degree ([Heckathorn, 2002](#); [Salganik and Heckathorn, 2004](#)). Weighting by the inverse of degree then gives estimates of population proportions; the most widely used version is the estimator of [Volz and Heckathorn \(2008\)](#). [Gile \(2011\)](#) proposed a successive sampling estimator that relaxes the assumption of sampling with replacement when the sample is a sizeable fraction of the population.

Methodological assessments have been sobering. Simulation on real networks showed that RDS estimates can be far more variable than early practice assumed ([Goel and Salganik, 2010](#)), and [Gile and Handcock \(2010\)](#) demonstrated sensitivity to seed selection, homophily and differential recruitment. Field practice has responded with formative research before sampling ([Johnston and Sabin, 2010](#)), with diagnostics that test the assumptions on the data collected ([Gile et al., 2015](#)) and with a reporting standard ([White et al., 2015](#)). These lessons carry over to any new population.

3. Why Lenders Are a Hard Case

The populations for which RDS was built, such as people who inject drugs, share an activity that brings them into regular contact and gives them little reason to exclude one another. Lenders differ in three ways that bear on the assumptions listed in Table 1.

First, lenders compete. In a market where clients and territory are contested, a lender may be unwilling to hand a coupon to a rival, and may not wish rivals to learn that he has spoken to researchers. Recruitment is then far from a random draw from the recruiter's contacts. Second, lending is often organised vertically. A financier may supply capital to several street-level lenders, who in turn employ collectors. Coupons can be expected to travel down such hierarchies, since a boss can direct subordinates to attend, and the resulting chains describe organisations more than peer networks.

Third, the population is heterogeneous and segmented. A study area may contain pawnbrokers who lend on the side, traders who advance goods on credit, salaried employees who lend to colleagues, organisers of savings clubs who also lend, and full-time professional lenders, some with links to

organised crime. These types may be separated by ethnicity, caste, gender or neighbourhood, with few ties across the divisions. Where sub-networks barely connect, chains started in one will not reach the others, and no weighting can repair the omission. The stages at which each difficulty arises are shown in Figure 1.

Table 1: Assumptions of respondent-driven sampling, how informal lending may violate them, and the proposed design response

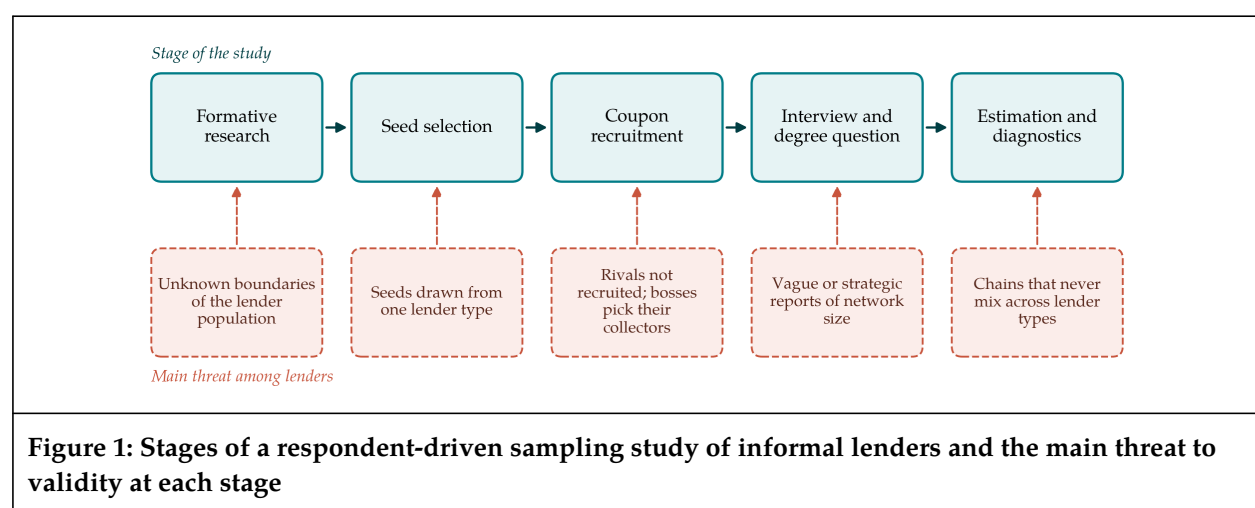
Assumption	Likely problem among lenders	Design response
Members know one another as members	Lending is concealed even from peers; side-line lenders are invisible	Formative work to map types; narrow, activity-based eligibility
Ties are reciprocal	Financier and collector know each other unequally	Record the relationship between recruiter and recruit
Network forms one connected component	Types divided by ethnicity, trade or territory	Seeds in every type; check for cross-type recruitment
Recruitment is random among ties	Rivals avoided; subordinates and kin favoured	Ask about coupon refusals and who was not approached
Degree is reported accurately	Boundary of "lender" unclear; incentive to understate	Specific, time-bounded question; sensitivity analysis
Sample is small relative to population	Local lender populations may be small	Successive sampling estimator; several sites

4. The Proposed Procedure

The procedure below follows the usual order of an RDS study, with adaptations for the problems just described.

4.1. Formative research and eligibility

Formative research is done before any coupon is issued. Through interviews with borrowers, community workers, traders and a few lenders willing to talk, the team maps the types of lender present, the terms they use for themselves, the places they can be met and the lines along which they are divided. This stage answers whether RDS is feasible at all. If lenders of different types do not know one another, the honest course is to treat them as separate populations with separate studies, or to use another method. Eligibility is then defined by activity and in local terms: for example, a person who has lent money at interest, to people outside the household, on a stated minimum number of occasions in a recent reference period within the study area. The wording should avoid legal categories that invite denial.



4.2. Seeds, coupons and incentives

Seeds are chosen purposively to cover every type and division identified, with more seeds than is usual in health applications because chains among lenders may die out early. Each respondent receives a small number of coupons; three is conventional, and fewer can be issued once recruitment is running well, to lengthen chains. Coupons carry a number and a contact for the study and nothing that identifies its subject. The incentive needs thought. A sum that motivates a street collector will mean little to a financier, and a sum large enough to interest a financier may attract impostors. Non-monetary incentives may serve better for wealthier lenders, such as a confidential summary of how local rates and practices compare, and screening questions on lending practice, asked by interviewers who know the trade, protect against people posing as lenders.

4.3. The interview and the degree question

Interviews are held at a neutral place or, since a fixed site may deter lenders, by appointment at a location the respondent chooses. No names are recorded, coupon numbers are the only link between records, and questions about specific borrowers, unlawful collection methods or links to named persons are left out. The degree question is asked in steps: how many people the respondent knows personally who lend in the study area, who also know the respondent, and with whom the respondent has spoken in a stated recent period. A further item asks how many of these the respondent would feel able to give a coupon to, which separates the network that exists from the network that can be recruited. The interviewer also records the relationship between respondent and recruiter (peer, employer, employee, relative, supplier of capital) and, at a follow-up contact, how many coupons were refused and by what kind of person.

4.4. Estimation and diagnostics

Estimates of proportions, such as the share of lenders who are full-time or who take collateral, would be computed with the Volz-Heckathorn and successive sampling estimators and reported side by side, with bootstrap intervals. Sample size planning should allow for a design effect well above one; [Salganik \(2006\)](#) suggested at least doubling the size a simple random sample would need, and later work implies that even this can be optimistic ([Goel and Salganik, 2010](#)). The diagnostics proposed by [Gile et al. \(2015\)](#) are applied as data arrive: plots of estimates against sample order to judge convergence, comparison of chains started from different seeds to detect bottlenecks, and tabulation of recruiter and recruit characteristics to measure homophily. The count of chains that cross between lender types is the single most informative figure. If there are none, estimates are reported for each type separately and no pooled figure is given.

5. Ethical and Legal Considerations

Where unlicensed lending is an offence, respondents are being asked about illegal activity, and the protection available to research records varies by jurisdiction. The team should establish before fieldwork whether its data could be demanded by police or regulators and should collect nothing that would identify a person if they were. Interviewer safety matters, particularly where lenders have ties to violent groups, and so does the safety of borrowers: a study must not hand lenders information about who has spoken to researchers. Peer recruitment also raises the risk of pressure when a financier recruits subordinates, which is one reason to record recruiter relationships and to make clear, privately, that declining carries no consequence. The reporting items of the STROBE-RDS statement ([White et al., 2015](#)) provide a checklist for documenting these decisions.

6. Limitations and What Remains to Be Shown

The procedure is untested, and there are reasons to doubt that RDS assumptions will hold among lenders. The proposal should therefore be read as a plan for finding out. A pilot in one bounded area would need to show four things: that seeds of each type recruit at all; that chains reach several waves and do not stall at the first; that some recruitment crosses between types and between organisations; and that reported degree is stable when the question is repeated in a different form. Failure on the first two would rule the method out. Failure on the third would confine it to describing types one at a time.

Even a successful pilot leaves limits. RDS estimates proportions within the networked population and says nothing about lenders who know no other lenders, a group that may be large among occasional side-line lenders. Estimating how many lenders there are is a separate problem; methods that use RDS data for population size exist (Handcock et al., 2014) but lean on further assumptions. And where comparison with an independent benchmark has been possible in other populations, RDS has not always recovered known values (McCreesh et al., 2012). For lenders no such benchmark exists, so findings should be compared with borrower-side surveys wherever they overlap.

7. Conclusion

Respondent-driven sampling offers a principled way to reach informal lenders without asking anyone to name names, and that alone makes it worth trying. Its inferential claims rest on assumptions that competition, hierarchy and segmentation among lenders put under strain. The design proposed here responds by investing heavily in formative research, seeding every lender type, measuring the recruitable network as well as the known one, recording how recruiter and recruit are related, and refusing to pool estimates across groups that the chains do not connect. Its worth can only be judged by a pilot, reported in enough detail for others to see where it failed.

Conflict of interest

To be declared by the author(s) before publication.

Funding

To be declared by the author(s) before publication.

References

- 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/10.1093/wber/4.3.329>
- Biernacki, P. and Waldorf, D. (1981). *Snowball sampling: Problems and techniques of chain referral sampling*. *Sociological Methods & Research*, 10(2): 141-163. <https://doi.org/10.1177/004912418101000205>
- Collins, D., Morduch, J., Rutherford, S. and Ruthven, O. (2009). *Portfolios of the Poor: How the World's Poor Live on \$2 a Day*. Princeton University Press, Princeton.
- Gile, K.J. (2011). *Improved inference for respondent-driven sampling data with application to HIV prevalence estimation*. *Journal of the American Statistical Association*, 106(493): 135-146.
<https://doi.org/10.1198/jasa.2011.ap09475>
- Gile, K.J. and Handcock, M.S. (2010). *Respondent-driven sampling: An assessment of current methodology*. *Sociological Methodology*, 40(1): 285-327.
<https://doi.org/10.1111/j.1467-9531.2010.01223.x>
- Gile, K.J., Johnston, L.G. and Salganik, M.J. (2015). *Diagnostics for respondent-driven sampling*. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 178(1): 241-269.
<https://doi.org/10.1111/rssa.12059>
- Goel, S. and Salganik, M.J. (2010). *Assessing respondent-driven sampling*. *Proceedings of the National Academy of Sciences*, 107(15): 6743-6747. <https://doi.org/10.1073/pnas.1000261107>
- Handcock, M.S., Gile, K.J. and Mar, C.M. (2014). *Estimating hidden population size using respondent-driven sampling data*. *Electronic Journal of Statistics*, 8(1): 1491-1521.
<https://doi.org/10.1214/14-EJS923>
- Heckathorn, D.D. (1997). *Respondent-driven sampling: A new approach to the study of hidden populations*. *Social Problems*, 44(2): 174-199. <https://doi.org/10.2307/3096941>
- Heckathorn, D.D. (2002). *Respondent-driven sampling II: Deriving valid population estimates from*

- chain-referral samples of hidden populations. *Social Problems*, 49(1): 11-34.
<https://doi.org/10.1525/sp.2002.49.1.11>
- Hoff, K. and Stiglitz, J.E. (1990). Introduction: Imperfect information and rural credit markets: Puzzles and policy perspectives. *The World Bank Economic Review*, 4(3): 235-250.
<https://doi.org/10.1093/wber/4.3.235>
- Johnston, L.G. and Sabin, K. (2010). Sampling hard-to-reach populations with respondent driven sampling. *Methodological Innovations Online*, 5(2): 38-48.
- McCreesh, N., Frost, S.D.W., Seeley, J., Katongole, J., Tarsh, M.N., Ndunguse, R., Jichi, F., Lunel, N.L., Maher, D., Johnston, L.G., Sonnenberg, P., Copas, A.J., Hayes, R.J. and White, R.G. (2012). Evaluation of respondent-driven sampling. *Epidemiology*, 23(1): 138-147.
<https://doi.org/10.1097/EDE.0b013e31823ac17c>
- Salganik, M.J. (2006). Variance estimation, design effects, and sample size calculations for respondent-driven sampling. *Journal of Urban Health*, 83(Suppl 1): 98-112.
<https://doi.org/10.1007/s11524-006-9106-x>
- Salganik, M.J. and Heckathorn, D.D. (2004). Sampling and estimation in hidden populations using respondent-driven sampling. *Sociological Methodology*, 34(1): 193-240.
<https://doi.org/10.1111/j.0081-1750.2004.00152.x>
- Venkatesh, S.A. (2006). *Off the Books: The Underground Economy of the Urban Poor*. Harvard University Press, Cambridge, MA.
- Volz, E. and Heckathorn, D.D. (2008). Probability based estimation theory for respondent driven sampling. *Journal of Official Statistics*, 24(1): 79-97.
- White, R.G., Hakim, A.J., Salganik, M.J., Spiller, M.W., Johnston, L.G., Kerr, L., Kendall, C., Drake, A., Wilson, D., Orroth, K., Egger, M. and Hladik, W. (2015). Strengthening the Reporting of Observational Studies in Epidemiology for respondent-driven sampling studies: "STROBE-RDS" statement. *Journal of Clinical Epidemiology*, 68(12): 1463-1471.
<https://doi.org/10.1016/j.jclinepi.2015.04.002>