Routledge
Taylor & Francis Group

Journal of Development Effectiveness

ISSN: 1943-9342 (Print) 1943-9407 (Online) Journal homepage: https://www.tandfonline.com/loi/rjde20

Demand driven replication research: an overview of financial services for the poor replication research

Benjamin DK. Wood, Abed Mutemi & Marie M. Gaarder

To cite this article: Benjamin DK. Wood, Abed Mutemi & Marie M. Gaarder (2019) Demand driven replication research: an overview of financial services for the poor replication research, Journal of Development Effectiveness, 11:4, 307-312, DOI: 10.1080/19439342.2019.1696871

To link to this article: https://doi.org/10.1080/19439342.2019.1696871

Published online: 29 Nov 2019.

JOURNAL OF DEVELOPMENT EFFECTIVENESS
2019, VOL. 11, NO. 4, 307–312
https://doi.org/10.1080/19439342.2019.1696871

Routledge
Taylor & Francis Group

ARTICLE

Demand driven replication research: an overview of financial services for the poor replication research

Benjamin DK. Wood[a], Abed Mutemi[b] and Marie M. Gaarder[c]

[a]Heifer International, Washington, DC, USA; [b]Amazon, WA, USA; [c]3ie, IN

ABSTRACT

Spurred on by the 'reproducibility crisis', social scientists are starting to adopt research transparency practices. Research funders are largely unaware that replication work could strengthen the reliability, rigour, and relevance of their investments. The Gates Foundation commissioned the International Initiative for Impact Evaluation (3ie) to award and quality assure the strongest financial services for the poor evidence. By working with the Gates Foundation to identify the studies, screen the applicants, and quality assure the seven replication research, 3ie ensured policy relevant papers. By publishing this special issue, 3ie is ensuring that the replication research is appreciated by the development community.

ARTICLE HISTORY
Received 7 November 2019
Accepted 20 November 2019

KEYWORDS
Replication; transparency; funding; financial services; development

Spurred on by the 'reproducibility crisis',[1] social scientists are starting to adopt practices that improve transparency in research processes, and increase confidence in study results. These practices range from pre-registering studies to providing access to study materials and data, to conducting replications.

Nevertheless, for many fields, research transparency continues to be more of a buzzword than an actuality. This is partly because the incentives and funding required to increase research transparency are largely missing. Journals continue to favour new research rather than replication work and to publish research whose data and codes have not been tested. Funders of research are largely unaware of the fact that various forms of replication work could strengthen the reliability, rigour and relevance of their investment. The International Initiative for Impact Evaluation (3ie) teamed up with the Gates Foundation to help change this dynamic. Through this partnership, the Gates Foundation commissioned 3ie to award and quality assure seven replication studies of the most policy relevant recent evidence around financial services for the poor in low- and middle-income countries.

This special issue is the next logical step in 3ie's push towards greater transparency.[2] By working with the Gates Foundation to identify the studies, screen the applicants, and quality assure the replication research, 3ie ensured that this demand driven replication research would be policy relevant. By publishing these replication studies in this special issue, 3ie is ensuring that the replication research will be acknowledged and appreciated by the larger development community.

Financial services for the poor

Every year, millions of people around the world transition out of poverty. Regional growth and economic opportunities like new jobs, technologies, and business opportunities, help people build

CONTACT Marie M. Gaarder mgaarder@3ieimpact.org

more stable economic lives. At the same time, millions of people remain trapped in a cycle of poverty that is very difficult to escape.

About 1.7 billion people worldwide are excluded from formal financial services (Global Findex, 2017), such as savings, payments, insurance, and credit. In developing economies, only 63 percent of adults have an account. Women are excluded from these beneficial financial systems more often. Nearly one billion are still left out of the formal financial system, and there is a 9 percent persistent gender gap in financial inclusion in developing economies.

Most poor households instead, operate almost entirely through a cash economy. This means they have to save in physical assets, such as livestock or jewellery. Cash gets spent, animals die, and jewellery can be lost or stolen. What's more, these forms of savings earn no interest and can actually lose value over time. To send money to family, those without a bank account have to rely on couriers or friends who carry cash by bus, which is expensive, insecure, and slow. To borrow money in an emergency, they must turn to moneylenders who charge notoriously high-interest rates.

Without formal financial histories, people are also cut off from potentially stabilising and uplifting opportunities like building credit or getting a loan to start a business. And it's harder to weather common financial setbacks, such as serious illness, a poor harvest, or an economic downturn. All too often, financial exclusion makes the expenses of poverty difficult to overcome.

The Bill & Melinda Gates Foundation's Financial Services for the Poor (FSP) program works to broaden the reach of low-cost digital financial services for the poor by supporting the most catalytic approaches to financial inclusion: helping to drive the development of digital payment systems that can help spread use of digital financial services quickly, advancing gender equality to ensure women share in the benefits of financial inclusion, and supporting the development of national and regional strategies that accelerate progress for the poor and provide exemplar models.

Motivation for funding the FSP replication window

FSP invests millions of dollars in driving financial inclusion across the world. In doing so, it relies heavily on research and evidence generated by the research community on how finance works and impacts the lives of poor people. Some of these research pieces although cited and referenced heavily have not been replicated or tested on the degree to which they generalise and apply in other contexts. This is one of the main reasons FSP partnered with 3ie: to review and replicate some of the research studies that have in one way or the other influenced the development community in allocating resources to interventions. The results of these replications can be used appropriately by FSP and other stakeholders to inform future investments.

Introducing the 6 papers

As mentioned above, lack of capital is a major concern in developing countries. Several advantages could be gained if farmers would have ready access to modern financial services allowing them, for example, to make optimal investment decisions or smooth consumption in case of sickness or poor harvests. An approach that goes beyond micro lending is to enable farmers to have access to formal saving mechanisms to build up individual capital. The study by Brune et al. (2016) is a field experiment based in Malawi, comparing different modes of formal yet more easily accessible saving mechanisms by cooperating with a local bank. In their analysis in this volume, Stage et al. consistently replicated the results and were able to extend the analysis by investigating the effect of specially offered savings vehicle. The replication study results show that this saving vehicle led to a substantial increase in farm output and profits. As the precise causal mechanism for this remains unclear, further research of this phenomenon would be a worthwhile endeavour.

The use of mobile phones as a payment mechanism is as a major innovation in financial services for the poor. The penetration and availability of mobile phones even in more remote areas are virtually unrivalled by established banks. The very low transactions costs of using mobile phones as

a mode of money transfers makes 'mobile money' a potentially promising tool for cash transfer programs. Hence, Aker et al. (2016) compared different modes of money transfer in Niger. Their study clearly demonstrated that 'mobile money' had a superior impact when compared to cash payments alone that cannot solely be explained by the (beneficial) effect of possessing a mobile phone. Beteta et al.'s replication study could fully confirm the findings of this analysis, while reporting that the magnitude of the beneficial effect may vary by age. Similarly, they could demonstrate that using mobile phones for transferring payments significantly reduces wasting (a disease caused by malnutrition) as compared to the other transfer modes.

M-PESA (M for mobile, PESA is Swahili for money) is a well-established Kenyan commercial financial service provider, who allows person-to-persons money transfers. This service is now widely used due to the large, relatively recent internal rural-to-urban migration movement that makes remittance a crucial source of income for poorer households. The study by Jack and Suri (2014) showed that the use of mobile money as a commercial financial service allows (better) informal risk sharing and improves overall household's financial health, i.e. households are less likely to reduce consumption and receive higher remittances when exposed to a financial shock. In this issue, Alinaghi replicated all the main findings and found them also robust to various robustness checks, in particular, alternative model specifications. In a further analysis, the existence of heterogeneity in treatment effects between urban and rural household was investigated, as rural households may be more likely to be underbanked, i.e. having less access to formal financial services compared to urban households. However, no significant heterogeneity when accounting for urban vs. rural context was identified.

Biometric smartcards, i.e. using electronic storage of money and finger scanners to identify the cardholder as the rightful recipient, are seen as an innovative form to allow the poor access to financial services and receive transfer payments in a transparent and safe manner reducing leakage of funds substantially. Muralidharan, Niehaus, and Sukhtankar (2016) conducted a large-scale randomised experiment investigating the effects of implementing such a biometric smartcard scheme for recipients of pensions and employment programs across eight districts within the Indian state of Andhra Pradesh. This experiment demonstrated that a biometric smartcard scheme increased employment and earnings of participants and reduced leakage of funds while ensured prompt access to transfer payments. Atanda's paper replicated the findings and found them robust when testing for outlies or variations of the model specifications. Moreover, Atanda investigated whether moderating factors can be identified, in particular, whether the treatment effects are heterogeneous across districts. This could not be confirmed, even though the coverage rate across districts was heterogeneous, varying from 31% to 100%.

In nutritional assistance programs traditionally transfers are made in-kind, e.g. handing out food parcels, to ensure that recipients consume enough of the 'right' goods. This approach is paternalistic and very costly to administer, in particular, compared to direct cash transfers. A middle ground is to disperse vouchers with a given cash value that can only be used in selected shops and for a pre-determined range of goods. Hidrobo et al. (2014) conducted a cluster randomised control trial in two provinces in Ecuador showing that all three modes of transfer (cash, food, vouchers) were superior to no assistance in terms of improving nutritional status, albeit to a varying degree. Lhachimi and Seuring could fully replicate the original findings while identifying that the magnitude of the effects varied by province, demonstrating the importance of considering the context in which an intervention takes place. They also extended the original cost-effectives-analysis by conducting a probabilistic sensitivity analysis, demonstrating that there is still substantial uncertainty which of these three modes of transfer is truly more cost-effective.

Unconditional cash transfers, i.e. without obligations, as a form of social assistance as compared to conditional cash transfer may have several advantages; in particular, they are less paternalistic, increase individual agency and have substantially lower administrative costs. The trial by Haushofer and Shapiro (2016) investigated the short-term impact of this intervention on economic outcomes and psychological well-being on poor households in Kenya showing improvements in several endpoints but not in

health, educations or female empowerment. In their replication study, Wang, Qiu, and Luo fully confirmed the original findings. Additionally, they investigated interaction effects at the village level. For several outcomes, in particular, psychological well-being, these interactions proved statistically significant, stressing the insight that the treatment effect also depends on contextual factors. Moreover, Wang et al. stress that long-term effects should be investigated more fully as a follow-up study by Haushofer and Shapiro (2018) that only improvement in asset holdings seem to be stable over time.

3ie's forward looking approach to replication

Research must meet the higher bar of being good enough for decision-making that affects human lives (not merely good enough for publication) – this is a tenet of 3ie, an organisation dedicated to evidence-informed decision-making in development. Therefore, 3ie is committed to continuously challenge the wider field of evaluators to improve the rigour of their research, and considers improvements in replication efforts a key component of this effort. 3ie considers that the future of replication will involve the following areas:

A recurring concern is that most replication work to date takes the arguments of the original authors at face value. Extensive attention is usually paid to the technical side of the analysis in the original article, while not enough attention is dedicated to results not reported (so reporting bias), the policy significance of the reported results, reflections about possible rival explanations for the results, and how the main variables were constructed. A more unbiased form of replication would be to allow the replication authors to see the original study write-up and results only *after* they have been asked to interpret the findings and draw conclusions.

Another replication deficiency in current research relates to the replication of qualitative research. While there is increasing acceptance that replication of quantitative research is part of best practice by funders and journals, replication in the qualitative research field is nascent. In a new initiative, 3ie is partnering with the Qualitative Data Repository (Syracuse University) to get their assistance in archiving and sharing select qualitative data, learn from the experience and thereby contribute to lessons and guidance on how to do this in the future. A further area in need of reflection and development is then to think about what type of replication work makes sense for what type of qualitative data, depending on the function of the qualitative data in the overall research. The equivalent of push button replication in the qualitative space, for example, may be an independent review of interview transcripts to verify that the same insights and conclusions are drawn. Field replication work could consist of, for example, focus group discussions being replicated following similar protocols in similar target populations. Sukumar and Metoyer (2019) provide a definition of what a replication of a qualitative study can potentially mean and entail, given that replication needs to be redefined for qualitative inquiry based on its inherent characteristics and interpretive strengths: *'A replication is a (novel) qualitative study conducted by independent researchers replicating one or more aspects (such as study design, research questions, context, methods, and participants) of an earlier qualitative study and embedding within its findings an interpretive comparison with a view to corroborate, elaborate, contrast, or clarify the elements corresponding to the replicated aspects with those of the earlier study.'*[3]

Finally, within the evidence architecture, 3ie has consistently promoted systematic reviews as the best and most unbiased source of policy-relevant evidence, and has carried out and supported the production of more than 50 systematic reviews over the last 10 years on various development interventions of global significance. In addition, 3ie hosts the largest repository of systematic reviews on international development issues.[4] Systematic reviews summarise the best available evidence on a specific question by synthesising the results of all relevant studies. The process is a systematic and transparent approach to all phases of a literature review. Another way of understanding systematic reviews is as a set of replications of studies in differentiated real-life settings. As discussed in Farrington et al. (2018), various framing conditions such as cultural, political, population, regional and time reduce the probability of obtaining consistent results from the same or similar

interventions or programs, *'therefore, it must be a core aim of replications to investigate those factors and processes that lead to variations in the findings of specific types of programs. This strategy would provide a realistic and differentiated guidance for practice and policy.'* Within the field of systematic reviews, 3ie is increasingly promoting an approach through which we plan the primary studies within a certain sub-sector ex ante in such a way that the ex post systematic review yields the most relevant policy insights (by, for example measuring similar outcomes in similar ways across studies, and by differentiating and subsequently controlling for particularly important framing conditions). Pure replication (desk replication) of systematic reviews is also a relatively nascent field, at least in the area of international development systematic reviews. At a minimum, 3ie is planning to introduce push button replication of any meta-analysis as a requirement before finalising any SR supported or produced by 3ie. In the medium term, we believe that search protocols can be replicated to verify that we end up with the same set of studies and subsequent results and conclusions.

Notes

1. Open Science Collaboration (2015).
2. 3ie has consistently served as a leader in the research transparency movement around impact evaluations in development. Over the 10 years of 3ie's existence, it has maintained a steadfast commitment to transparency, including requiring data sharing for 3ie commissioned research (Goel, Khatua, and Gaarder 2019). From creating a replication program (Brown, Cameron, and Wood 2014) to expanding the program to incorporate push button replication (Wood, Müller, and Brown 2018), 3ie staff defined, implemented, and published replication research. Currently, 3ie's research transparency policy (3ie 2018), including the push button replication of all final reports and the requiring of open data and code within 6 months of completing the research, is on the cutting edge of transparency requirements across development organisations.
3. Sukumar and Metoyer (2019), p.2.
4. http://www.3ieimpact.org/evidence-hub/publications/systematic-reviews.

Acknowledgments

The authors would like to thank valuable contributions from Stefan Lhachimi and Emmanuel Jimenez in conceptualising and encouraging this paper. We would also like to thank all of the replication researchers and original authors who made these replication studies possible. All errors are our own.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

We greatly appreciate funding from the Bill and Melinda Gates Foundation that faciliated the replication studies in this speical issue.

Notes on contributors

Benjamin DK. Wood, Heifer International, formerly International Initiative for Impact Evaluation, 3ie.

Abed Mutemi, Amazon, formerly Bill & Melinda Gates Foundation, Financial Services for the Poor.

Marie M. Gaarder, International Initiative for Impact Evaluation, 3ie.

ORCID

Benjamin DK. Wood http://orcid.org/0000-0002-3942-4025

References

3ie. 2018. "3ie." *Research Transparency Policy*. https://www.3ieimpact.org/sites/default/files/2019-06/3ie-research-transparency-policy.pdf

Aker, J. C., R. Boumnijel, A. McClelland, and N. Tierney. 2016. "Payment Mechanisms and Antipoverty Programs: Evidence from a Mobile Money Cash Transfer Experiment in Niger." *Economic Development and Cultural Change* 65 (1): 1–37. doi:10.1086/687578.

Brown, A., D. Cameron, and B. Wood. 2014. "Quality Evidence for Policymaking: I'll Believe It When I See the Replication." *Journal of Development Effectiveness* 6 (3): 215–235. doi:10.1080/19439342.2014.944555.

Brune, L., X. Giné, J. Goldberg, and D. Yang. 2016. "Facilitating Savings for Agriculture: Field Experimental Evidence from Malawi." *Economic Development and Cultural Change* 64 (2): 187–220. doi:10.1086/684014.

Collaboration, O. S. 2015. "Estimating the Reproducibility of Psychological Science." *Science* 349 (6251): aac4716. doi:10.1126/science.aac4716.

Farrington, D. P., F. Lösel, R. F. Boruch, D. C. Gottfredson, L. Mazerolle, L. W. Sherman, and D. Weisburd. 2018. "Advancing Knowledge about Replication in Criminology." *Journal of Experimental Criminology* 15 (3): 373–396.

Findex, G. 2017. "World Bank." https://globalfindex.worldbank.org/

Goel, N., S. Khatua, and M. Gaarder. 2019. "Making Data Reusable: Lessons Learned from Replications of Impact Evaluations." *Plos One*. Forthcoming. doi:10.1371/journal.pone.0222701.

Haushofer, J., and J. Shapiro. 2016. "The Short-term Impact of Unconditional Cash Transfers to the Poor: Experimental Evidence from Kenya." *The Quarterly Journal of Economics* 131 (4): 1973–2042. doi:10.1093/qje/qjw025.

Haushofer, J., and J. Shapiro. 2018. "The Long-Term Impact of Unconditional Cash Transfers: Experimental Evidence from Kenya." *Working Paper*, Princeton University.

Hidrobo, M., J. Hoddinott, A. Peterman, A. Margolies, and V. Moreira. 2014. "Cash, Food, or Vouchers? Evidence from a Randomized Experiment in Northern Ecuador." *Journal of Development Economics* 107: 144–156. doi:10.1016/j.jdeveco.2013.11.009.

Jack, W., and T. Suri. 2014. "Risk Sharing and Transactions Costs: Evidence from Kenya's Mobile Money Revolution." *American Economic Review* 104 (1): 183–223. doi:10.1257/aer.104.1.183.

Muralidharan, K., P. Niehaus, and S. Sukhtankar. 2016. "Building State Capacity: Evidence from Biometric Smartcards in India." *American Economic Review* 106 (10): 2895–2929. doi:10.1257/aer.20141346.

Sukumar, P. T., and R. Metoyer. 2019. "Replication and Transparency of Qualitative Research from a Constructivist Perspective." doi:10.31219/osf.io/6efvp.

Wood, B., R. Müller, and A. Brown. 2018. "Push Button Replication: Is Impact Evaluation Evidence for International Development Verifiable?" *PLoS ONE* 13 (12): e0209416. doi:10.1371/journal.pone.0209416.