Microplots and food security: encouraging replication studies of policy relevant research

Benjamin Douglas Kuflick Wood and Maria Vasquez

Abstract

Replication research is a valuable, yet often misunderstood, tool for increasing our understanding of promising research findings. In their short paper below the authors discuss their principles for conducting replication research, explain how they picked a candidate study for replication, describe the robustness checks they would conduct in their replication plan, and give an overview of how they would interpret their replication results. The authors also describe some of their lessons learned after working in replication research for over five years.

(Published in Special Issue The practice of replication)

Authors
Benjamin Douglas Kuflick Wood, International Initiative for Impact Evaluation, Washington Office, USA, ben.wood@heifer.org
Maria Vasquez, Independent Researcher

Citation Benjamin Douglas Kuflick Wood and Maria Vasquez (2017). Microplots and food security: encouraging replication studies of policy relevant research. Economics Discussion Papers, No 2017-71, Kiel Institute for the World Economy. http://www.economics-ejournal.org/economics/discussionpapers/2017-71

I.   A general discussion of principles about how one should do a replication

We developed our replication philosophy after both working on replication research with the International Initiative for Impact Evaluation (3ie). When conducting internal replication, where we reproduce the published results using the existing data, we would start by drafting a replication plan. These plans, similar to pre-analysis plans, outline what we intend to examine within their study. By pre-specifying the replication robustness checks, we maintain transparency around what we originally intended to check in their study. It also forces us to justify any additions we might include in the study after posting the replication plan.

We would then follow 3ie's standard replication process by conducting these four stages to our replication study: i) a push button replication (PBR), ii) pure replication, iii) measurement and estimation analysis, and iv) theory of change analysis. In the first two steps, we verify the original results, using the original research methods. Then in the last two steps we assess the strength of the original results to pre-specified robustness checks (Brown *et al.* 2014). Given the sensitivities around replication studies, we will avoid using terminology like "error" or "mistake" during our replication process (Brown and Wood, 2014).

We would start our study by verifying the ability to reproduce the published results by conducting a push button replication. This verification process begins by pre-specifying the key results that the replication will compare (Wood et al. 2017). Then we attempt to run the original code on the original data. We would then compare our PBR results to the original results, with pre-specified decision rules indicating if any differences found classify as "major" or "minor."

Push button replication (PBR) is the stage where we are more comfortable assessing the direct replicability of the original publication, as we will follow a clear protocol, with results classified as *comparable, minor differences, major differences, etc.* (Wood et al. 2017). Classifying these results, while still contingent on thresholds of difference, is less controversial than other forms of replication in that PBR removes interpretation from the replication process.

We would then assess the reproducibility of the paper by recoding the original results, the pure replication stage. Using the same data and same methodology, as described in the publication and supporting material, we would independently reproduce the original findings.

Finally, we would conduct the robustness exercises we described in our replication plan for the measurement and estimation analysis, and the theory of change analysis stages. We apply the replication diagnostic from Brown and Wood (2017) to determine the most appropriate robustness checks for our replication study. We conduct these robustness exercise to assess if we, as independent researchers, are able to produce similar results when using plausible analysis strategies to address the same research question with the same data as in the original publication.

By conducting our replication study following a replication plan and certain guidelines, we are trying to move beyond discussing the need for more replication research in the social sciences. While numerous calls have highlighted a desire for more replication studies, we oftentimes find a general incentive incompatibility problem and a lack of clarity on how to conduct replication

research prevents researchers from conducting these studies (see Duvendack and Palmer-Jones (2013) for an example of the replication incentives discussion). We hope this study, along with all the papers in this special issue, will encourage researchers to participate in this crucial step within the scientific process.

II.   An explanation of why the "candidate" paper was selected for replication

Santos *et al.*'s "Can government-allocated land contribute to food security? Intrahousehold analysis of West Bengal's microplot allocation program" (2014) assessed the impact of the Nijo Griha, Nijo Bhum (NGNB) program, which provides small land plots to landless poor households. These microplots are titled in the name of female household members or jointly for some couples. The researchers expected the program to affect a set of intermediate outcomes: tenure security, agricultural investments, use of credit for agriculture, and women's participation in household decisions. In turn, they expected long-term outcomes to eventually reduce hunger vulnerability, while increasing protein consumption, distribution of food within the household, and dietary diversity.

The article was selected for replication because it contributes to the understanding of an intervention that tackles food insecurity, a pressing development issue. The Food and Agriculture Organization (FAO) estimates that in the period from 2010-2012 the prevalence of undernourishment in the developing regions was 12.9% (FAO 2015). Ensuring access to land and security of tenure have been recognized as important components of food security policies (HLPE 2013). Considering the potential high productivity of small farms and their impact on food security, programs that provide access to microplots may significantly improve household welfare. It has been suggested that providing access to small plots, even fractions of an acre, creates a household safety net while contributing to improved household nutrition and income. In peri-urban and urban areas, several evaluations have found some indications of a positive association between these small agriculture interventions and food security, although the evidence base is weak, and there is a need for more rigorous research (see Warren *et al.* 2015; Poulsen *et al.* 2015). In that line, Santos et al. provide information about the effectiveness of the intervention, contributing to the literature on securing land rights and agriculture investment (Deininger 2006; Goldstain & Udry 2008; Besley 1995), women empowerment through land security and children's health (Allendorf 2007), and homestead agriculture for improving nutrition (Talukder *et al.* 2010).

Finally, given its encouraging results on intermediate outcomes, the study is also relevant because it can have direct policy implications for the design of programs that expand access to land, especially in India. This type of program has already been implemented in different parts of India, including West Bengal, Karnataka, Andhra Pradesh and Odisha. The paper can further inform the debate in India regarding the bill that entitles landless rural households to access plots of 0.1 acre (Government of India 2013), which has been on hold since 2013 (Draboo 2015). As various levels of the Government of India consider scaling up this program, testing the robustness of the results to reevaluation will help policymakers better evaluate the effectiveness of this intervention.

III.   A replication plan that applies these principles to the "candidate" article

Our plan to conduct the replication includes four stages that range in complexity from simply reproducing the results with the original code and data to analyzing the theory of change of the program. The first two stages intend to examine if the results are reproducible, or if they can be replicated based only on the paper. In the third and fourth stages of the replication, we will our analysis on the intermediate results and the impact of the program following its theory of change.

In the third and fourth stages we use the replication diagnostic to focus our robustness checks of the original publication on the validity of assumptions, data transformations, estimation methods, and heterogeneous impacts (Brown and Wood 2017). Since the main results of the paper are at the intermediate level, our robustness checks will reexamine the influence of providing a microplot on household: i) perceived land tenure security, ii) likelihood of access credit for agriculture, iii) use of improved inputs, and iv) women's likelihood to be involved in important food and agriculture decisions. If possible, we will also explore impact level effects of the program on food security. This analysis will allow us to investigate if these results are sensitive to varying hunger measurements and the methodologies used in the original publication.

We will focus on the land tenure security results, as the original authors find statistically significant increases in program participants that tracked to increases in plot sizes. Plot size requirements are highly relevant to the research question, both because of cost and land distribution constraints. In particular, we will undertake an exploratory analysis to determine if a plot size threshold exists, under which microplots are too small to influence household outcomes.

We also plan to conduct additional tests not reported by the original authors. We will examine if the intervention affected household wealth, which we find plausible given the intermediate results reported. Furthermore, we will independently assess the results for households headed by women.

1. Push button replication

As explained in the first section, the first step of the replication research is using the original code and data to attempt to replicate the results of the paper. To do so, we use the same software, data, and code that was used by the authors. The paper identifies positive impacts in a range of intermediate outcomes related to i) perceived land tenure security, ii) likelihood of access credit for agriculture, iii) use of improved inputs, and iv) women's likelihood to be involved in important food and agriculture decisions. In each area the authors present several outcomes that are closely related to each other. Considering the theory of change, we classify the following intermediate results as key results for our PBR: i) Woman reports that her household will have same or more access and control over the plot in five years, ii) Household has taken out a loan from a bank since 2009 iii) Share of household land over which woman respondent decides "How to use the plot", iv) Household used seedlings, seeds, or grafted stems in last year. The original authors present these results in table 3 of the original paper.

4

After running the code on the data, we will assess if the results are *comparable*, or if they have *minor* or *major* differences with the published paper.

2. Pure replication

A second stage is to conduct a pure replication analysis. Following Wood and Dong (2015), we will recode the analysis to reproduce the original paper, using the original data and following the methodology presented in the paper and in the supporting documents. Then, we will compare the original code with the recreated code and we will analyze any differences. We will focus on identifying differences on the management of outliers and data imputation. We will also pay special attention to examine the inverse propensity score score-weighted regression technique used in the original analysis to estimate the causal effect, as it is not discussed at length in the paper.

3. Measurement and Estimation Analysis (MEA)

We will focus our measurement and estimation analysis on the validity of the original research assumptions, data transformations, and estimation methods. Within research assumptions, we will include attritor households in the analysis that did not have females in the household at the time of follow-up. When looking at data transformations, we will examine different ways to measure hunger vulnerability and convert children into adult equivalency units. When considering estimation methods, our robustness checks will develop a timeline of the intervention's implementation and, if feasible, conduct an annual analysis of the intervention, test the robustness of the results to district level estimations, and conduct a difference-in-difference and a treatment on the treated analysis. We will outline the reasoning behind each of these robustness checks below.

3.1 Research assumptions: attrition

The original authors report a fairly large 25 per cent attrition rate between baseline and follow-up survey. We plan to re-examine the original authors' attrition analysis and test the robustness of the intermediate and impact level results to inclusions of attriters, if information was collected on households without women present. If information on these households were not collected, we will further explore the possibility of differential attrition.

3.2 Data transformations: outliers, data imputation and variable construction

Published articles usually provide little documentation around outlier identification and missing data imputation. We will examine decisions on data transformation to test the robustness of the results to these decisions. If missing values or outliers prove to be a factor in the original

analysis, we will test the robustness of the results to inclusion of observations dropped by the original analysis or alternative imputation techniques.

As this paper focuses on household hunger, we plan to test the robustness of the impacts on food security to an alternative measurement of hunger vulnerability. The original research uses a binary proxy indicator, assuming hungry households experienced times within the last three

5

months when they did not have food and/or money to purchase food. Due to the length of the recall period and the stark contrast between the two options, we plan to explore the use of categorical variables to capture more variation within household hunger vulnerability. With these measurements, we expect to be able to observe if there were changes in hunger vulnerability that were not captured by the original analysis.

In addition to general household hunger vulnerability, the original authors explore heterogeneous hunger impacts based on gender and age. We plan to look at alternative ways to account for age. The authors define three categories of people: adults who are 12 years and older, children who fall between the ages of 4 and 11 and infants who are aged of 0-3. Without a clear explanation for creating the adult threshold at age 12, we will use standard adult equivalency units to convert younger people into adults and test the robustness of the overall hunger results.

3.3 Estimation methods: type and level of analysis

The timeline of the intervention, as presented in the paper, is difficult for us to follow. We will develop an intervention and survey timeline. We plan to clarify with the original authors when the treatment households knew they had access to the land, when the baseline surveys were conducted, as the original publication notes "data collection for our baseline took place in two rounds: May-June of 2010 and January-June of 2011" (Santos *et al.* 2014 p. 864). If baseline information is specific for individual households, we will test the robustness of the original results to an annual analysis that will assess the effects for households that had access to the land for a longer period. If households that owned the land for longer periods show improved outcomes, this would further support the theory of change of the paper.

The original authors evaluated the pilot intervention while it was rolled out in three districts. As noted in the original authors' balance tables, statistically significant imbalances exist between treatment and control in two of these three areas. We plan to test the robustness of the original results to district level analyses of the outcomes of interest.

We further plan to test the robustness of the intermediate results to a difference-in-difference analysis. It appears that the original authors use the baseline data to conduct the matching exercise, and then compare the differences between the treatment and control households in the follow-up period. Given the existence of baseline data, assuming that similar questions were asked in the baseline survey instrument, we anticipate testing for shifts in outcomes of interest based on household inclusion in the treatment group.

Santos et al. find a number of their results statistically insignificant at the intent to treat (ITT) level. As only around 25 per cent of the treatment households actually relocated to their new microplot at the time of follow-up, and it is unclear how many other treatment households actually received land, we will calculate treatment on the treated (TOT) estimates. The original authors conduct these estimates for the food security outcomes. We will extend their analyses to the intermediate outcomes, which are the main focus of the publication findings.

4. Theory of Change Analysis

In our theory of change analysis we propose to look at three topics that fall within the heterogeneous impacts category (Brown and Wood, 2017). First we want to examine alternative outcomes of interest, specifically with respect to wealth outcomes. Next, we plan to test the effectiveness of the intervention within the subsample of single/divorced/widowed women. Finally, focusing more on the policy relevance of these results, we will explore intermediate outcomes above and below 5 and 10 decile levels.

4.1 Heterogeneous impacts: outcomes of interest and sub-samples

The original publication tests a number of self-observed wealth measures, both in terms of intermediate and final outcomes. The researchers note in their theory of change that they expect to see increases in land investment, possibly through access to more credit. Increased agricultural related income from land ownership represents another potential channel of increased investment, which may also lead to greater household food security. We plan to test for possible changes in household wealth due to microplot ownership if surveys collected this information from households.

We will test the alternative hypothesis that the intervention mainly influences female headed households, instead of changing power dynamics in married relationships. Based on the summary statistics, 17 per cent of the treatment households are composed of single, divorced, or widowed women. If sample limitations prevent us from conducting a full heterogeneous impact analysis, we will look at how closely the outcomes of these women correlate with the general sample.

In the middle of the project the Government of West Bengal decided to cap land distribution at 5 deciles per household. The original researchers note that the greater the microplot land allocation, the larger the intermediate outcome food security effects. We would like to take this one step further by focusing specifically on households that received 5 or less deciles of land and 10 or less deciles of land. We will test the robustness of the results by focusing on these sub-samples of the survey population. This analysis is important from a cost-effectiveness perspective, as the program costs increase with larger government land purchases and allocations.

IV. A discussion of how to interpret the results of the replication

Providing landless households with microplots is touted as an economically feasible and politically appealing intervention in a number of countries. In Bangladesh, Hillenbrand and Waid (2014) discuss the importance of microplots in increasing micronutrient levels and general household food security. More generally, the Bureau for Development Policy in the United Nations Development Programme (UNDP) recommends microplots as a method to reduce inequality in the developing world (UNDP, 2013).

To interpret the findings of our replication study, we will focus on the robustness of the intermediate results to our measurement and estimation re-analysis. We focus on these results,

as the original authors use them to demonstrate the policy relevance of providing microplots to landless households. On one hand, the paper's results would be weakened if is it shown that they are not robust to the difference-in-difference analysis or microplot threshold for land tenure security was economically unfeasible. On the other hand, the results would be strengthened if: i) it is shown that the effects are stronger for households that had the microplots longer time (annual analysis), and ii) if the intermediate results are stronger for households that actually received land (TOT analysis). Additionally, since the paper does not find impacts on food security, the case for microplots would be strengthened if the measurement and estimation analysis shows that the program had effects on food security by: i) examining minimum land thresholds, and ii) including alternative measurement of hunger vulnerability.

Replication studies should be a tool for broader discussion of influential and innovative research. Therefore, our interpretation of the results of the replication process does not focus on arguing a "success" or "failure" to replicate the original study, instead we aim to deepen the research dialogue by simply reporting our findings and how do they strengthen or weaken the main takeaways of the original article. We will detail our attempt to independently verify these policy relevant results, and we will then invite the development community as a whole to review, assess, and comment on the original publication and the replication report. We hope this concrete example contributes to the use of replication research as a tool to encourage a greater conversation about what development programs worked well and what evidence looks promising for scaling.

Bibliography

Allendorf, K, 2007. Do women's land rights promote empowerment and child health in Nepal? *World Development*, 35, pp. 1975–88. Available at: http://www.sciencedirect.com/science/article/pii/S0305750X07001416 [Accessed 5 May 2017].

Besley, T, 1995. Property rights and investment incentives: theory and evidence from Ghana. *The Journal of Political Economy*, 103, pp. 903–937. Available at: http://www.journals.uchicago.edu/doi/10.1086/262008 [Accessed 5 May 2017].

Brown, A, Cameron, D and Wood, B, 2014. Quality evidence for policymaking:The article I'll believe it when I see the replication. *Journal of Development Effectiveness*, 6(3), pp. 215-235. Available at: http://www.tandfonline.com/doi/abs/10.1080/19439342.2014.944555 [Accessed 5 May 2017].

Brown, A and Wood, B, 2014. When is an error not an error? Guest post by Annette N. Brown and Benjamin D. K. Wood [online]. Available at: https://blogs.worldbank.org/impactevaluations/when-error-not-error-guest-post-annette-n-brown-and-benjamin-d-k-wood [Accessed 6 April 2017].

Brown, A and Wood, B, 2017. Which tests not witch hunts: A diagnostic tool for conducting replication research, under review.

Collier, P and Dercon, S, 2014. African agriculture in 50 years: Smallholders in a rapidly changing world? *World Development*, 63, pp. 92–101. Available at: http://www.sciencedirect.com/science/article/pii/S0305750X13002131 [Accessed 6 April 2017].

Deininger, K and Songqing, J, 2006. Tenure security and land-related investment: Evidence from Ethiopia. *European Economic Review*, 50(5), pp. 1245-1277. Available at: http://www.sciencedirect.com/science/article/pii/S0014292105000577 [Accessed 5 May 2017].

Draboo, A, 2015. Why the Modi government must work on land reform before land acquisition [online]. Available at: http://www.landesa.org/why-the-modi-government-must-work-on-land-reform-before-land-acquisition/ [Accessed 10 September 2015].

Duvendack, M and Palmer-Jones, R, 2013. Replication of quantitative work in development studies: Experiences and suggestions. *Progress in Development Studies*, 13(4), pp. 307-322. Available at: http://journals.sagepub.com/doi/abs/10.1177/1464993413490480?journalCode=pdja [Accessed 5 May 2017].

FAO, IFAD and WFP, 2015. The State of Food Insecurity in the World 2015. Meeting the 2015 international hunger targets: taking stock of uneven progress. Rome: FAO. Available at: http://www.fao.org/3/a-i4646e.pdf [Accessed 5 May 2017].

Goldstein, M and Udry, C, 2008. The profits of power: land rights and agricultural investment in Ghana. *Journal of Political Economy*, 116 (6), pp. 981–1022. Available at http://www.journals.uchicago.edu/doi/abs/10.1086/595561 [Accessed 5 May 2017].

Government of India, 2013. *The National Right to Homestead Bill, 2013.* Available at: https://canvas.uw.edu/courses/882885/files/26051447/download?verifier=CYu5A3y5zoSbIbNyiQ4r09nOEkq8mWaDpxEHexnp [Accessed 5 May 2017].

Hillendbard, E and Waid, J, 2014. Adapting food-based strategies to improve the nutrition of the landless: A review of HKI's homestead food production program in Bangladesh, a chapter in Thompson, B and Amoroso, L Improving diets and nutrition: Food-based approaches. Rome, Italty: Food and Agricultural Organization.

HLPE, 2013. *Investing in smallholder agriculture for food security. A report by the High Level Panel of Experts on Food Security and Nutrition of the Committee on World Food Security.* Rome, Italty: HLPE. Available at: http://www.fao.org/fileadmin/user_upload/hlpe/hlpe_documents/HLPE_Reports/HLPE-Report-6_Investing_in_smallholder_agriculture.pdf [Accessed 5 May 2017].

Poulsen, M, McNabb, P, Claytonb, M and Neffc, R, 2015. A systematic review of urban agriculture and food security impacts in low-income countries. *Food Policy*, 55, pp. 131–146. Available at: http://www.sciencedirect.com/science/article/pii/S0306919215000809 [Accessed 5 May 2017].

Santos, F, Fletschner, D, Savath, V and Peterman, A, 2014. Can government-allocated land contribute to food security? Intrahousehold analysis of West Bengal's microplot allocation program. *World Development* 64, pp. 860-872. Availble at: http://www.sciencedirect.com/science/article/pii/S0305750X14002265 [Accessed 5 May 2017].

Talukder, A, Haselow, NJ, Osei, AK, Villate, E, Reario, D, Kroeun, H, SokHoing, L, Uddin, A, Dhungel, S and Quinn,V, 2010. Homestead food production model contributes to improved household food security and nutrition status of young children and women in poor populations - lessons learned from scaling-up programs in Asia (Bangladesh, Cambodia, Nepal and Philippines). *Field Actions Science Reports* [online], Available at*: http://factsreports.revues.org/404* [Accessed 11 September 2015].

UNDP, 2013. Humanity divided: Confronting inequality in developing countries.New York, USA: UNDP. Available at: http://www.undp.org/content/dam/undp/library/Poverty%20Reduction/Inclusive%20development/Humanity%20Divided/HumanityDivided_Full-Report.pdf [Accessed 31 August 2017].

Warren, E, Hawkesworth, S and Knai, C, 2015. Investigating the association between urban agriculture and food security, dietary diversity, and nutritional status: a systematic literature review. *Food Policy*, 53, pp. 54–66. Available at: http://www.sciencedirect.com/science/article/pii/S030691921500024X [Accessed 5 May 2017].

Wood, B and Dong, M, 2015. Recalling extra data: a replication study of Finding missing markets, 3ie Replication Paper 5. Washington, DC: International Initiative for Impact Evaluation (3ie). Available at: http://www.3ieimpact.org/media/filer_public/2015/05/19/rps5-recalling_extra_data_Nl8blco.pdf [Accessed 5 May 2017].

Wood, B, Brown, A, Djimeu, E, Vasquez, M, Yoon, S and Burke, J, 2017. Replication protocol for push button replication [online]. Available at: https://osf.io/2juhv/ [Accessed 7 April 2017].

Please note:

You are most sincerely encouraged to participate in the open assessment of this discussion paper. You can do so by either recommending the paper or by posting your comments.

Please go to:

http://www.economics-ejournal.org/economics/discussionpapers/2017-71

The Editor