Replication of a Research Claim from Cohen et al. (2015),
from *American Economic Review*

Replication Team: Melba Tutor and Esteban Méndez-Chacón

Research Scientist: Anna Abatayo

Action Editor: Miguel Fonseca

Independent Reviewers
(add name below when you initiate review, comment “DONE” on your name when you finish):

Reviewer #1: Eirik Strømland
Reviewer #2: [NAME]

Reviewer #3: [NAME]

Review Period: September 14 - September 21

View-only links to: Original Paper, Original Materials, Replication Data, Replication Analysis

Privacy Statement: Other teams are making predictions about the outcomes of many different studies, not knowing which studies have been selected for replication. As a consequence, the success of this project requires full confidentiality of this peer review process. This includes privacy about which studies have been selected for replication and all aspects of the discussion about these replication designs.

Instructions for Data Analysts

The preregistration for this replication study was started by a separate team of researchers who were responsible for identifying data sources and constructing them into a replication dataset(s) for your use in the analysis. They have completed sections 1-13 of the preregistration below, and included additional materials in the OSF project that document how the dataset was constructed.

In cases where all of the underlying data sources were able to be freely shared and posted, the constructed dataset(s) have been posted to the OSF as well, which you are free to use in designing the analysis plan (see below for details). In cases where some or all of the data sources could *not* be freely shared or posted, the replication dataset(s) are not provided on the OSF. Rather, you will need to follow the instructions and code to first reconstruct the datasets, and then proceed with your work. In such cases, the team responsible for creating the dataset(s) has provided summary statistics in the OSF that correspond to the constructed datasets, so you can verify that the datasets you create match what they intended.

You'll be responsible for filling out sections 16-25 of the preregistration below. Before you do so, **please review the original study, sections 1-15 of the preregistration, and the materials provided on the OSF**, so that you are familiar with all of the decisions that have been made to date. In many cases, the 'data preparer' will have left you instructions and suggestions on how the provided data can be used in the analysis, as well as idiosyncrasies and discrepancies in the data that you should be aware of. The data preparers have tried to be thorough in including all variables that you might need, but please keep in mind the following:

- Some of the variables included in the constructed dataset(s) may not be needed in the final analysis, so please do not feel the need to necessarily use all of the provided variables.
- Some of the variables needed might have mistakenly been excluded from the constructed datasets. If you find that this is the case, please let Andrew or Anna know, and they will work with you to supplement the datasets as needed.

For these secondary data replications, we would like the analysis plan to be completed before the preregistration goes through review, so that after review, the only remaining steps are registration and running the analysis code on the full datasets. To facilitate that, we are asking that you include in section 19 a link to the code you will use that takes the constructed dataset(s) provided to you and produces the focal analysis (including all of the cleaning, merging, and transforming required). When developing your analysis plan and code, please randomly sample 5% of the data for use in your work and demonstrate that the focal analysis produces sensible results using just that random sample by providing a screenshot of the output (see section 19 for details). **Do not use the rest of the data until after your study is registered and it is time to run the final analysis**. In section 19, you will find a statement that we are asking you to bold that confirms you've only used 5% of the data when developing and testing your code. If this approach will not work for any reason, please let Andrew or Anna know and disclose deviations from this plan somewhere in the preregistration.

- In cases where we are providing you a complete dataset, you can just sample out 5% of the observations and hold the rest out until you are ready to perform the final analysis.
- In cases where we are providing you multiple datasets that need to be combined prior to analysis, please sample out 5% of the observations in whatever way is most sensible.
    - For example, in cases where each dataset contains complete observations on its own (a typical 'row bind' situation), it makes the most sense to sample out 5% of each dataset separately and then combine them together to develop and test your code.

    - In cases where datasets need to be merged in order to create complete observations (a typical 'column bind' situation), it makes the most sense to merge the separate datasets into a full dataset first, and then sample out the 5% before proceeding with the rest of the analysis code.
- We leave the decision on how to sample out the random subset of data to you, so long as (a) you are not performing any analyses on the complete dataset until after your study is registered and (b) whatever decision you make is documented in the preregistration.

Finally, in cases where the replication data combines observations from the original study with observations that were not used in the original study (what we are calling 'hybrid replications'), please perform up to three analyses (details immediately below). This will likely require you to subset your data, based on the description of the original analysis provided in the study.

- When the 'new' data alone can clear the minimum power threshold, please perform one analysis that relies only on the 'new data' (the focal analysis), one analysis that relies on all available data, and a third analysis that relies only on the original data. Please make sure all three analyses are documented (with code) in section 19 below.
- When the 'new' data alone *cannot* clear the minimum power threshold, please perform one analysis that combines all available data, and a second that only uses the old data. Please make sure both analyses are documented (with code) in section 19 below.

Please contact Andrew or Anna if you have any questions. After you've completed the remaining sections of the preregistration and uploaded all the necessary materials to the OSF, please contact the SCORE coordinators regarding next steps.

Preregistration of Cohen_AmEcoRev_2015_2lb5

Existing Data Replication

Study Information

1. Title (provided by SCORE)

RR TEAM INSTRUCTIONS: *This has been determined by SCORE**.***

Replication of a research claim from Cohen et al. (2015) in *American Economic Review*.

2. Authors and affiliations

RR TEAM INSTRUCTIONS: *Fill in the names and affiliations of your team below**.***

Melba Tutor [Data identification and preparation][1]
Esteban Méndez-Chacón[2]

1 Not applicable
2 Central Bank of Costa Rica

3. Description of study (provided by SCORE)

RR TEAM INSTRUCTIONS: *This description has been provided by SCORE. Please review and make a SCORE project coordinator aware of any edits, additions, and corrections you would suggest to the paragraph. You are free to add additional descriptions of your project in a separate paragraph.*

The claim selected for replication from Cohen et al. (2015) is that all three subsidy levels lead to a large and significant increase in ACT (artemisinin combination therapies) access. This reflects the following statement from the paper's abstract: "We show that a very high subsidy (such as the one under consideration by the international community) dramatically increases access, but nearly one-half of subsidized pills go to patients with-out malaria." The authors study impacts on ACT access (as well as other measures of treatment-seeking behavior) by presenting results from regression equation (2). The dependent variable is "Took ACT". The focal independent variable is "Any ACT subsidy". Panel A of Table 2 presents a specification where all three ACT subsidies are pooled and compare outcomes to the control group. Column 1 of Table 2 reports results on overall ACT access. The coefficient for "Any ACT subsidy" is 0.187 with robust standard errors clustered at the household level of 0.038, significant at the 1 percent level.

4. Hypotheses (provided by SCORE with possible Data Analyst additions)

RR TEAM INSTRUCTIONS: *The focal test for SCORE is indicated as H*. If you will test additional hypotheses (or use alternate analyses) that help you to evaluate the claim your replication/reproduction is testing, number them H1, H2, H3 etc. (You can place H* in the list wherever makes sense). Please make sure that any additional hypotheses are logical deductions/operationalizations of the selected SCORE claim or are necessary to properly interpret the focal H* hypothesis. Research that is outside this scope should be described in a separate preregistration.*

Specific points to keep in mind (please also consult the Reviewer Criteria):

- *Are the listed hypotheses specific, concise, clearly testable, and specified at the level of operationalized variables?*
- *Are hypotheses identified as directional or non-directional, and, if applicable, have the direction of hypotheses been stated? (Example: "Customers' mean choice satisfaction will be higher in the CvSS architecture condition than in the standard attribute-by-attribute architecture condition.")*
- *Does the list of hypotheses/tests indicate whether additional hypotheses are taken from the original study or modified/added by the team?*

H*: ACT [artemisinin combination therapies] subsidies induce take-up of ACT.

In this replication attempt, the focal hypothesis can be operationalized as follows: diagnosis-dependent ACT subsidies induce take-up of ACT. The replication data is sourced from a study where the intervention is a combination of a free diagnostic test and an ACT voucher conditional on testing positive. Further details are provided in Section 8.

Design Plan

5. Study type

NOTE: *The study type selected should be based on the data collected for the replication, and not necessarily the data used in the original study.*

- **Experiment - A researcher randomly assigns treatments to study subjects, this includes field or lab experiments. This is also known as an intervention experiment and includes randomized controlled trials.**
- Observational Study - Data is collected from study subjects that are not randomly assigned to a treatment. This includes surveys, natural experiments, and regression discontinuity designs.
- Meta-Analysis - A systematic review of published studies.
- Other

[QUESTION 5 - BOLD YOUR RESPONSE ABOVE]

6. Blinding

RR TEAM INSTRUCTIONS: *Select any/all of the below that apply for your study by bolding them. You will give a longer description in the next question.*

- No blinding is involved in this study.
- For studies that involve human subjects, they will not know the treatment group to which they have been assigned.
- Personnel who interact directly with the study subjects (either human or non-human subjects) will not be aware of the assigned treatments. (Commonly known as “double blind”)
- **Personnel who analyze the data collected from the study are not aware of the treatment applied to any given group.**

[QUESTION 6 - BOLD YOUR RESPONSE ABOVE]

7. Blinding

RR TEAM INSTRUCTIONS: *Since all existing data replications are based on data that has already been collected, in most cases it will not be necessary to comment on participant blinding. In the rare instance when an existing experiment is being re-analyzed for an existing data replication and blinding is a relevant consideration, please provide below any details regarding blinding that are important for a reviewer to be aware of.*

In the replication data source, data analysts were blinded from the treatment assignment of study areas until all the survey waves have been conducted. Due to the nature of the intervention — diagnostic testing and ACT vouchers —, study participants cannot be blinded from their treatment assignment.

8. Study Design

RR TEAM INSTRUCTIONS: *Please describe how data was collected in the original study and how it compares to the data that was selected for the replication attempt. Explain why the data selected for the replication study is suitable for a replication and if any substantial deviations exist between the two.*

If the data used in the replication combines observations from the original study with new observations (e.g. if the data selected for the replication attempt comes from the same longitudinal survey as the original study), describe how ‘original’ and ‘new’ observations relate to each other and an estimate for what proportion of the final dataset’s observations will be comprised of original vs. new observations.

Specific points to keep in mind (please also consult the Reviewer Criteria):

- *Does the preregistration specify the unit of analysis?*
- *Does the preregistration provide sufficient detail about how the data selected for the replication attempt deviates from or is congruent with the data employed in the original study?*
- *Does the preregistration describe whether and how 'original' and 'new observations' are combined together for the replication dataset?*

The Cohen (2015) study conducted a randomized control trial in 3 districts of Kenya where malaria is endemic. They selected four drug shops in four rural market centers and sampled all households in the catchment area (within a 4-kilometer radius) of each of these shops. (p.619) There are 2 core treatment arms: (1) vouchers for ACT subsidy and (2) vouchers for ACT and rapid diagnostic test (RDT) subsidies. Within each treatment arm, subjects were also distributed into 3 subsidy levels. The control group received no subsidy.

The study period is from May to December 2009. A baseline and an endline survey were conducted, tracking the same households. The endline survey was conducted four months after the vouchers were distributed. The data used for the focal analysis is from the subsample of households that received any ACT subsidy only. The unit of analysis is the first illness episode with at least one malaria-like symptom that the household experienced following the baseline. (p. 627)

The data for the replication attempt is from a randomized controlled trial in 32 community clusters in western Kenya. These study areas were selected based on availability of retail outlets selling ACTs and the presence of a community health worker (CHW) program. In intervention areas, CHWs perform malaria RDTs for any individual 1 year old and above experiencing a malaria-like illness. Malaria RDT-positive individuals that get tested through the CHWs receive a voucher for subsidized ACT.

The study period is from July 2015 to May 2017. The data for the replication attempt consist of 4 independent cross-sectional household surveys, conducted at baseline, and at 6, 12, and 20 months post baseline. In each cross-sectional survey, households in which at least one member above 1 year old had a fever or malaria-like illness in the last 1 month were sampled through systematic random sampling. The selected random household and sampling interval were different in each survey to minimize sampling the same household or febrile individual in multiple waves. Only one random fever per sampled household was selected for the analysis.

The data selected for the replication attempt is suitable for several reasons: (1) it also employed an experimental design; (2) its study areas are in Western Kenya but in different districts; (3) its study period is 6 years after the original study; and (4) the subsidy levels in the original and replication data are both priced vis-à-vis the target subsidy levels of the Affordable Medicines Facility for malaria (AMFm) program.

The main deviation of the replication data from the original study is that the intervention is a combination of a free diagnostic test and an ACT voucher conditional on testing positive. The

focal analysis is the effect of any ACT subsidy on ACT take-up, regardless of whether the illness is malaria, and regardless of whether it was tested. The replication data cannot separate the effect of the ACT subsidy from the free and positive RDT test. Nonetheless, the data selected for the replication attempt can have a valuable contribution to the appreciation of the focal analysis as it focuses on a relevant question from a public health perspective, i.e. do individuals who have information on their malaria status get induced to buy medicines through subsidies.

9. Randomization (free response)

RR TEAM INSTRUCTIONS: *If the variables used for this replication attempt were randomized, state how they were randomized, and at what level.*

The unit of randomization in the replication data is the community cluster, which is an administrative unit with around 1,000 households. There are 32 community clusters enlisted for the study and these were randomly assigned 1:1 to control and intervention arms. Thus, the treatment or the focal independent variable is randomized.

Sampling Plan

This section describes how the data sources for the replication were selected, how they were prepared into a replication dataset, and the number of observations that will be analyzed from these data. Please keep in mind that the data described in this section are the actual data used for analysis, so if you are using a subset of a larger dataset, please describe the subset that will actually be used in your study.

10. Existing data

1.1.1. Registration prior to creation of data
1.1.2. Registration prior to any human observation of the data
1.1.3. Registration prior to accessing the data
1.1.4. Registration prior to analysis of the data
1.1.5. Registration following analysis of the data

[QUESTION 10 - BOLD YOUR RESPONSE ABOVE]

11. Explanation of existing data

NOTE: *For replications that rely on existing data sources, this question refers to the data that will be used for the replication analysis (i.e. the final replication dataset), and not (a) the data from the original study or (b) the data sources accessed to construct the replication dataset. Since no new data will be created for 'existing data replications,' 1.1.1 should never be selected. Since all analyses will occur after registration, 1.1.5 should also never be selected.*

The data for this replication attempt has been downloaded and prepared for replication prior to registration. Variables were selected based on congruence with the variables used in the original study, as determined from the description in the codebook or the journal articles. All the observations from the raw replication data (baseline, and 3 independent cross-sectional surveys post-baseline) are included in this replication attempt.

12. Data collection procedures

RR TEAM INSTRUCTIONS: *Please describe the process for constructing the replication dataset in as much detail as you can. The sections below should be used to provide the following information:*

- *Which variables are needed from the original study to perform a good-faith, high-quality replication.*
- *Which data sources were used, why they were selected, any deviations between the original study design and the replication study design that these selections present, and the procedures used to access the data.*
- *Which of the variables from the original study are available in the replication data sources, including relevant details about each measure.*
- *The procedure for creating the replication dataset, in both narrative and script form.*
- *A data dictionary that documents each variable included in the replication dataset.*

In the sections below, please provide links to the original materials whenever possible -- including descriptions of the original datasets and corresponding codebooks. If materials can be shared on the OSF, please do so, and provide view-only links to those materials.

Specific points to keep in mind for reviewers:

- *Does the preregistration describe which data sources were selected for the replication study and why each is suitable?*
- *Does the preregistration make clear how the data sources were used to construct the replication dataset?*

(a) Data Needed

RR TEAM INSTRUCTIONS: *List below the datasets and variables the original author used to analyze the focal claim. Include details regarding the sample size, waves or years used, and other details pertinent to finding an existing dataset for replication. Please include page numbers when excerpting from the original article. If possible, categorize the list of variables as one of the following: dependent variable, focal independent variable, control variable, or sample parameters/clustering variable. Finally, include the sample size of the original study's focal analysis, if it is available.*

The sample size of the focal analysis is 631 observations, which consist of the first illness episode with at least one malaria-like symptom that the household experienced after the baseline. The unique number of households is 575.

Dependent Variable(s)

Take-up of ACT

- =1 if the illness episode is treated with ACT

Focal Independent Variable(s)

Any ACT subsidy

- =1 if the household received any of the three ACT subsidies

Control Variable(s)

Age of the household head

- There are two variables used for the age of the household head: (1) B_head_age_imputed, which is the age of head at baseline, with missing replaced by sample mean; and (2) B_head_age_missing, which is =1 if the age of the household head at baseline is missing.

Sample Parameters

Strata

- Stratum fixed effects (n=28)

Cluster

- Household ID (n=575)

(b) Data Access

RR TEAM INSTRUCTIONS: *Describe below the data sources that will provide the replication variables. Include information such as the name of the data source (e.g., Indonesian Family Life Survey), the description and link of the data source, and the waves needed to create a final replication dataset.*

Also describe the process for accessing the data sources that will be used to create the final replication dataset; specify how long long it took for the registration to be approved and what information was required (e.g., writeup of the purpose of the project, email address from an IPCSR institution, etc.); and verify that the data can be opened as expected. If applicable, provide a link to the page where you registered to access the data.

Describe in detail any restrictions on data access and data-sharing, as well as any additional terms of data use that will be relevant for the replication study and final report (e.g. citations that will need to be made). If you were able to access the data because of special permissions that you have, but that you expect other researchers might not have, please document those as well.

The data for this replication attempt is from the study "Improving rational use of ACTs through diagnosis-dependent subsidies: evidence from a cluster-randomized controlled trial in western Kenya" by Wendy Prudhomme O'Meara, et.al (2018) in PLOS Medicine. The data can be accessed from the Dryad Digital Repository using this link. From this link, the following files can be downloaded automatically (no registration needed): (1) main dataset; (2) data dictionary. Both files are in MS Excel formats.

The data publication file states that "This work is licensed under a CC0 1.0 Universal (CC0 1.0) Public Domain Dedication license. This releases your work to the public domain for any use." The recommended citation for the data is as follows: Prudhomme O'Meara, Wendy et al. (2019), Data from: Improving rational use of ACTs through diagnosis-dependent subsidies: evidence from a cluster-randomized controlled trial in western Kenya, Dryad, Dataset, https://doi.org/10.5061/dryad.59p4111.

The full study protocol for the randomized trial is also published and publicly available. The citation for this protocol is as follows: Laktabai J, Lesser A, Platt A, et al. Innovative public–private partnership to target subsidised antimalarials: a study protocol for a cluster randomised controlled trial to evaluate a community intervention in Western Kenya. BMJ Open 2017;7: e013972. doi:10.1136/bmjopen-2016-013972.

(c) Variable Availability

RR TEAM INSTRUCTIONS: *For each variable required for the replication analysis (listed above), describe the variables from the replication data that can be used to measure it (including which data files or sources each measure is found in),* ***any notes a data analyst should consider when using the measure in a replication analysis****, and any important differences between the original variable and the proposed replication variable.*

If there are multiple variables in the replication data that correspond to a required variable (e.g. two different measures of education in the replication data), include all of those options below. If a variable from the original study ***cannot*** *be measured using the replication data, please make that clear as well.* ***Finally, include a description of the identifiers used to merge multiple datasets, if applicable.***

All the variables listed below are in ReplicationData_Cohen_AmEcoRev_2015_2lb5.dta.

Dependent variable: Take-up of ACT

Variable name: drugs_taken_AL

- Description: =1 if the subject took the ACT Artemether Lumefantrine for the current illness
- Additional notes: The ACT voucher used in the original study was Coartem, which is an Artemether Lumefantrine type of ACT. Thus, this is the equivalent outcome variable.

Focal independent variable: Any ACT subsidy

Variable name: maltest_chw_voucher_given

- Description: =1 if the subject was given an ACT voucher subsidy
- Additional notes: By study design, the ACT voucher subsidy is offered to individuals that tested positive in a rapid diagnostic test administered by the community health worker.

Control variable: Age

Variable name: patient_age_categ

- Description: Patient age is divided into 3 categories: code 0 = under 5 years old; code 1 = 5 to 17 years old; code 2 = 18 years old and above.
- Additional notes: Only the categories of age of the eligible individual is available. The age of the household head, which is the variable used in the original study, is not available in the replication data. The original study included the household head's age as a control variable because there is a lack of balance across treatment groups in the age composition of households. In particular, the control group has older household heads (page 624).

(d) Data Creation

RR TEAM INSTRUCTIONS: *Create a dataset using the data sources and variables listed above. Provide a detailed narrative describing how the various datasets were cleaned and merged into a final replication dataset. Provide a view-only link to a clearly commented script on the OSF that produces the replication data as described in the narrative. Our preference is that this be either an R script or a script from another language that similarly allows for open and reproducible analyses. Please let the SCORE team know if this is not possible.*

- *If the data can be freely shared and posted to OSF, please post it in your OSF project and provide a link to the completed dataset below.*
- *If any part of the dataset cannot be shared between researchers or posted to the OSF, please leave the final dataset off the OSF. Instead, include either below or in your script (commented out at the bottom) two pieces of information that will help an independent team verify they have created the dataset according to your instructions:*
    - *The dimensions of the final dataset(s) you've created (# of rows, # of columns)*
    - *A summary of 8-10 variables in the replication dataset. For numeric variables, the summary should include the mean, standard deviation, and count of NAs. For categorical variables, the summary should include each level present in the data*

and its count, as well as a count of NAs. If multiple datasets are submitted as part of your work, at least one variable should be included from each dataset.

The data from the replication sources should be preserved in as 'raw' a form as possible, in order to give the data analyst the most latitude to clean the variables as they see fit. Variables from the original source should be preserved in their original form (e.g. do not recode values of 99 to NA). New variables should only be created when they're needed to complete the merge or combine the datasets; in those cases, please preserve a version of the original, unaltered variable in the new dataset.

Please also use this section to describe:

- *Any deviations between the original study design and the replication design that would result from using this replication dataset.*
- *Any notes about using these variables that you would like to pass along to the data analyst.*

The raw replication data is preserved. Thus, the main step in creating the replication data is to import the raw data into Stata, apply the variable labels, and save it in Stata format. The script for this process is here and the data file that results is here.

(e) Data Dictionary

RR TEAM INSTRUCTIONS: *Create a data dictionary following this template. Provide below a view-only link to the completed data dictionary included in the OSF project. If the Data Analyst will need to create new variables using the variables in the final replication dataset (e.g. recoding the provided education variable to be in a better format for analysis), please document below your recommendation on how the analyst should do so. Please also document any additional notes regarding the variables in the dataset that do not fit within the provided data dictionary template or the other sections above.*

The data dictionary for the replication data is found here. All the variable names in the raw data are preserved. The variable labels and description were revised for clarity based on the source data dictionary and dataset. Aside from the information required in the template, the source data dictionary contains the following: whether the variable is derived or not, and the question text, values, and notes per survey wave. Pertinent notes that might be of help to the data analyst are already added in the data dictionary created for this replication attempt.

13. Sample size

RR TEAM INSTRUCTIONS: *Please report below the analytic sample size(s) in the replication dataset, with reference to however many units or levels are in the data. Please report as much information here as will be helpful for the review committee to be aware of, including differences in sample size resulting from various analytic decisions (e.g. listwise deletion vs multiple imputation).* ***Finally, when the replication combines observations from the original study***

with new observations, please estimate what proportion of the analytic sample's observations will be comprised of original vs. new observations.

Data finders' response goes here: The replication data has a total of 7,416 observations from the baseline and 3 post-baseline cross-sectional surveys. The total observations from the 3 post-baseline surveys is 5,399. Each observation in each wave corresponds to a unique household, i.e. only 1 random fever is selected in each sampled household. However, it is possible that the same household or febrile individual may have been surveyed more than once across the survey waves. There is no way to determine this in the data, but the data source reports that this is "expected to be rare" because the random starting household and the sampling interval are both different in each wave.

The focal dependent and independent variables have 5,177 and 504 non-missing observations from the 3 post-baseline surveys, respectively. By design, the ACT subsidy is given to individuals who were tested through the CHW and got a positive result. Individuals in treatment areas who were not tested, or who were tested in government clinics or other channels are not eligible for the ACT subsidy.

Required sample size [to be filled out by the SCORE team]: The primary unit of analysis is the household. An estimate of the minimum viable sample size for the data analytic replication is: 94. For comparison, the stage1 required sample size would be: 446 and the stage2 sample size would be: 1000.

Notes: From the table legend, the guess for this power analysis is that in most cases it's 1 observation per household, but that it is not certain. The power analysis assumes the original and replication have the same data structure in terms of the same number of observations per household. Based on Table 2, there are 631 observations in the analysis, where observation is, 'first illness episode with at least one malaria-like symptom that the household experienced following baseline. A few households have multiple observations if multiple household members were ill simultaneously'. We could not find the number of households in this particular analysis in the paper, but based on the output from the push button replication attempt of this study, we believe the number of households is 575, meaning an average of 1.1 observations per household. The replication assumes that the original and replication have the same data structure, and so to meet the minimum threshold the replication would need 104 observations in total and at least 94 households. If the replication gets to the minimum number of observations by having more households but only 1 observation per household, it will have more power than expected.

14. Sample size rationale

For data analytic replications in SCORE, three sample sizes are calculated:

- *A minimum threshold sample size, defined as the sample size required for 50% power of 100% of the original effect*
- *A stage 1 sample size, defined as the sample size needed to have 90% power to detect 75% of the original effect*
- *A stage 2 sample size, defined as the sample size needed to have 90% power to detect 50% of the original effect*

Details about how those sample sizes were calculated for this project are found here: https://osf.io/bgh9j/?view_only=95960c6821ef4d3d9e733b2ee101b04b

15. Stopping rule (provided by SCORE)

None of the observations from the replication data overlap with the original data. Accordingly, the SCORE team recommends that a single analysis be performed that incorporates all available observations.

Variables

RR TEAM INSTRUCTIONS: *The preregistration form divides variables across three questions: manipulated variables, measured variables, and indices (i.e. analytic variables derived from raw variables). For existing data replications, only fill out the "Measured variables' and 'Indices' sections. Please do not fill out anything in the 'Manipulated variables' section.*

The raw data of any transformed variable (e.g. reaction time → log reaction time) or any created index should be defined in the 'Measured variables' section. Details regarding the variable transformation should be specified in the 'Transformations' section. Details regarding the creation of an index should be specified in the 'Indices' section.

Across these questions, you should define all variables that will later be used during your analysis (including data preparation/processing). You can describe all variables in the preregistration and/or summarize and link to a data dictionary (codebook) in your repository to answer these questions.

If you will share data from your replication, this is also the place to state whether any variables will be removed prior to sharing the dataset (e.g. to reduce risk of participant identification or comply with copyright restrictions on scale items.)

16. Manipulated variables

RR TEAM INSTRUCTIONS: *Manipulated variables in this preregistration refer specifically to variables that have been randomly assigned in an experiment. The use of data from an experiment should be rare in existing data replications. If your existing data replication relies on experimental data, please document each manipulated variable as a measured variable, and use the codebook to indicate what each level of the variable corresponds to (e.g. participants assigned to the treatment condition = 1; participants assigned to the control condition = 0). The default language in bold below has been copied into all existing data replication preregistrations.*

N/A -- not documented for existing data replications.

17. Measured variables

RR TEAM INSTRUCTIONS: *Please use this section to document each variable that was used in the original study's analysis and the role it served (e.g. dependent variable, control variable, sample parameter, etc). For each variable, provide the description of the variable offered in the paper and/or codebook of the original study, the variable in the replication dataset that it corresponds to, and explain any deviations between the two. In cases where an equivalent replication variable was not found, explain how, if at all, you expect it will affect the replication attempt. In cases where you are adding a variable that was not present in the original study, please explicitly state that you are doing so, and explain how, if at all, you expect it will affect the replication attempt.*

Specific points to keep in mind (please also consult the Reviewer Criteria):

- *Does the preregistration surface all of the variables needed to replicate the focal analysis?*
- *Are deviations between the original variables and replication variables documented when needed?*

Note: Lines 18 to 255 of the Stata do-file "Cohen et al 2015 - Replication Analysis 5% random sample.do" generate the variables described below.

TOOK_ACT

- Use in the analysis: Dependent variable.
- Description from the original study: "y_{eh} is the outcome of interest for illness episode e in household h" (page 627) and "Column 1 of Table 2 reports results on overall ACT access" (page 628)
- Variables used in the replication: DRUGS_TAKEN_AL.
- TOOK_ACT corresponds to a dummy variable equal to 1 if the subject took the ACT Artemether Lumefantrine for the current illness, and 0 otherwise.

- No deviations between the original study and the replication study.

ACT_SUBSIDY

- Use in the analysis: Focal independent variable.
- Description from the original study: “Panel A of Table 2 presents a specification where we pool all three ACT subsidies and compare outcomes to the control group, while panel B presents a specification where we separately estimate the impact of the three different subsidy levels. In both cases, the omitted category is the “no ACT subsidy” (control) group” (page 627)
- Variables used in the replication: MALTEST_CHW_VOUCHER_GIVEN.
- ACT_SUBSIDY corresponds to a dummy variable equal to 1 if the subject was given an ACT voucher subsidy, and 0 otherwise. Note that as in Cohen et al (2015), the omitted category is the control group (no ACT subsidy).
- For cases where MALTEST_CHW_VOUCHER_GIVEN=98 (“Don’t remember”), ACT_SUBSIDY is considered a missing value.

HH_ID

- Use in the analysis: Identification variable.
- Description from the original study: “Robust standard errors clustered at the household
- level in parentheses” (page 627)
- HH_ID assigns an unique number to each household in the replication sample.
- Note: According to the "Sample Size" section, each observation in each wave corresponds to a unique household, i.e. only 1 random fever is selected in each sampled household. However, it is possible that the same household or febrile individual may have been surveyed more than once across the survey waves. There is no way to determine this in the data, but the data source reports that this is “expected to be rare” because the random starting household and the sampling interval are both different in each wave. Consequently, it is assumed that each observation is a different household.

WEIGHT

- Use in the analysis: Sampling weight.
- Description from the original study: The analysis should take into account the nature of the data, by including sampling weights in the analysis.
- Variables used in the replication: WEIGHT.
- WEIGHT is a continuous variable that indicates the sampling weights.

MALTEST_WHERE

- Use in the analysis: Sample parameter variable.
- Description from the original study: As it is explained in the “Study Design” and "Sample size" sections, by design, the ACT subsidy is given to individuals who were tested through the community health worker (CHW) program and got a positive result. Consequently, the estimate is restricted to the subsample where MALTEST_WHERE=1 (CHW)

- Variables used in the replication: MALTEST_WHERE.
- MALTEST_WHERE is a variable that indicates where a malaria test was taken.

WAVE

- Use in the analysis: Sample parameter variable.
- Description from the original study: As it is explained in the "Sample size" section, the focal dependent and independent variables are obtained from the 3 post-baseline surveys, i.e., WAGE!=0.
- Variables used in the replication: WAVE.
- WAVEis a variable that indicates the data collection period.

STRATA

- Use in the analysis: Control variable.
- Description from the original study: "$_{strata}$ are strata fixed effects" (page 627)
- Variables used in the replication: CU_CODE.
- STRATA is a numerical variable that indicates the cluster unit, the unit of randomization.

Table 1 in Cohen et al (2015) presents baseline household characteristics and tests for balance across treatment groups. To test balance across the experimental groups, they regressed each dependent variable in Table 1 on a dummy variable for each of the three ACT subsidy levels and a dummy variable for the RDT subsidy. Moreover, they include a full set of strata dummies in the regression. They find no significant differences across treatment groups, other than for the number of acres owned and the age distribution in the household. In particular, because the control group has slightly older household heads, and age is highly correlated with malaria positivity, they decided to control for the age of the household head to avoid any confounding in their results (pages 622 to 624).

Following this experimental design, the control variables that are included in the replication are the ones that are significantly different between the treatment and the control groups, at a 5% level. **For this, the full replication dataset is used.** The baseline household characteristics that are compared across the groups are:

1. SES_HH_ITEMS. Description: Household assets. It includes the variables:
    a. ELECTRICITY: a dummy variable equal to 1 if the household has electricity, and 0 otherwise. (SES_HH_ITEMS=1). It is considered as missing if SES_HH_ITEMS is missing.
    b. TELEVISION: a dummy variable equal to 1 if the household has a television, and 0 otherwise. (SES_HH_ITEMS=2). It is considered as missing if SES_HH_ITEMS is missing.
    c. REFRIGERATOR: a dummy variable equal to 1 if the household has a refrigerator, and 0 otherwise. (SES_HH_ITEMS=3). It is considered as missing if SES_HH_ITEMS is missing.

d. RADIO: a dummy variable equal to 1 if the household has a radio, and 0 otherwise. (SES_HH_ITEMS=4). It is considered as missing if SES_HH_ITEMS is missing.
e. MOBILE: a dummy variable equal to 1 if the household has a mobile phone, and 0 otherwise. (SES_HH_ITEMS=5). It is considered as missing if SES_HH_ITEMS is missing.
f. MOTORCYCLE: a dummy variable equal to 1 if the household has a motorcycle, and 0 otherwise. (SES_HH_ITEMS=6). It is considered as missing if SES_HH_ITEMS is missing.
g. CAR/TRUCK: a dummy variable equal to 1 if the household has a car/truck, and 0 otherwise. (SES_HH_ITEMS=7). It is considered as missing if SES_HH_ITEMS is missing.
h. BANK_ACCOUNT: a dummy variable equal to 1 if the household has a bank account, and 0 otherwise. (SES_HH_ITEMS=8). It is considered as missing if SES_HH_ITEMS is missing.
i. NO_ASSETS: a dummy variable equal to 1 if the household has no assets, and 0 otherwise. (SES_HH_ITEMS=9). It is considered as missing if SES_HH_ITEMS is missing.

Table 1 presents the results of regressing each variable above on the dummy variable ACT_SUBSIDY. The results suggest that households in the treatment group are significantly more likely to have a refrigerator and a mobile phone, at a 5% significance level. Consequently, REFRIGERATOR and MOBILE are considered as part of the control variables in the replication.

Table 1 - Baseline Summary Statistics for Household Assets (SES_HH_ITEMS)

	Control group mean (1)	ACT_SUBSIDY (2)
ELECTRICITY	0.526	0.003
	(0.093)***	(0.081)
TELEVISION	0.524	0.005
	(0.093)***	(0.027)
REFRIGERATOR	0.526	0.107
	(0.093)***	(0.030)***
RADIO	0.495	0.019
	(0.095)***	(0.014)

MOBILE	0.369	0.041
	(0.098)***	(0.011)***
MOTORCYCLE	0.527	-0.009
	(0.092)***	(0.011)
CAR/TRUCK	0.526	-0.016
	(0.093)***	(0.022)
BANK_ACCOUNT	0.520	0.010
	(0.094)***	(0.005)
NO_ASSETS	0.526	-0.112
	(0.093)***	(0.177)

Following Table 1 in Cohen et al (2015, page 623), the first column shows average values of characteristics for the control group. Column 2 shows regression coefficients and standard errors on the treatment group. All regressions include a full set of strata dummies.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.

2. SES_TOILET_TYPE. Description: Type of toilet. It includes the variables:
    a. FLUSH_TOILET: a dummy variable equal to 1 if the household has a flush or pour flush toilet, and 0 otherwise. (SES_TOILET_TYPE=1). It is considered as missing if SES_TOILET_TYPE is missing.
    b. VIP_TOILET: a dummy variable equal to 1 if the household has a VIP/Ventilated improved pit, and 0 otherwise. (SES_TOILET_TYPE=2). It is considered as missing if SES_TOILET_TYPE is missing.
    c. LATRINE_WITH_SLAB: a dummy variable equal to 1 if the household has a pit latrine with slab, and 0 otherwise. (SES_TOILET_TYPE=3). It is considered as missing if SES_TOILET_TYPE is missing.
    d. LATRINE_WITHOUT_SLAB: a dummy variable equal to 1 if the household has a pit latrine without slab, and 0 otherwise. (SES_TOILET_TYPE=4). It is considered as missing if SES_TOILET_TYPE is missing.
    e. COMPOSTING_TOILET: a dummy variable equal to 1 if the household has a composting toilet, and 0 otherwise. (SES_TOILET_TYPE=5). It is considered as missing if SES_TOILET_TYPE is missing.

f. BUCKET_TOILET: a dummy variable equal to 1 if the household has a bucket toilet, and 0 otherwise. (SES_TOILET_TYPE=6). It is considered as missing if SES_TOILET_TYPE is missing.
g. NO_FACILITY_TOILET: a dummy variable equal to 1 if the household has no facility/bush/field, and 0 otherwise. (SES_TOILET_TYPE=7). It is considered as missing if SES_TOILET_TYPE is missing.
h. OTHER_TOILET: a dummy variable equal to 1 if the household has another type of toilet, and 0 otherwise. (SES_TOILET_TYPE=8). It is considered as missing if SES_TOILET_TYPE is missing.

Table 2 presents the results of regressing each variable above on the dummy variable ACT_SUBSIDY. Note that no household in the treatment or control groups has a flush toilet or a bucket toilet. Consequently the variables FLUSH_TOILET and BUCKET_TOILET are not included in Table 2.

The results suggest that households in the treatment group are significantly more likely to have a VIP/Ventilated improved pit, a composting toilet, and another type of toilet, at a 5% significance level. Consequently, VIP_TOILET, COMPOSTING_TOILET, and OTHER_TOILET are considered as part of the control variables in the replication.

Table 2 - Baseline Summary Statistics for Type of Toilet (SES_TOILET_TYPE)		
	Control group mean (1)	ACT_SUBSIDY (2)
VIP_TOILET	0.526	0.353
	(0.093)***	(0.078)***
LATRINE_WITH_SLAB	0.533	-0.033
	(0.094)***	(0.055)
LATRINE_WITHOUT_SLAB	0.503	0.030
	(0.101)***	(0.054)
COMPOSTING_TOILET	0.526	0.476
	(0.093)***	(0.114)***
NO_FACILITY_TOILET	0.530	0.172
	(0.094)***	(0.247)
OTHER_TOILET	0.526	0.195

	(0.093)***	(0.062)***
Following Table 1 in Cohen et al (2015, page 623), the first column shows average values of characteristics for the control group. Column 2 shows regression coefficients and standard errors on the treatment group. All regressions include a full set of strata dummies. *** Significant at the 1 percent level. ** Significant at the 5 percent level.		

3. SES_FLOOR_MATERIAL. Description: Type of floor material. It includes the variables:
    a. EARTHEN_FLOOR: a dummy variable equal to 1 if the household has a earthen floor, and 0 otherwise. (SES_FLOOR_MATERIAL=1). It is considered as missing if SES_FLOOR_MATERIAL is missing.
    b. CEMENT_FLOOR: a dummy variable equal to 1 if the household has a cement floor, and 0 otherwise. (SES_FLOOR_MATERIAL=2). It is considered as missing if SES_FLOOR_MATERIAL is missing.
    c. FLOOR_TILES: a dummy variable equal to 1 if the household has floor tiles, and 0 otherwise. (SES_FLOOR_MATERIAL=3). It is considered as missing if SES_FLOOR_MATERIAL is missing.
    d. WOOD_PLANKS: a dummy variable equal to 1 if the household has wood planks, and 0 otherwise. (SES_FLOOR_MATERIAL=4). It is considered as missing if SES_FLOOR_MATERIAL is missing.
    e. POLISHED_WOOD: a dummy variable equal to 1 if the household has a polished wood floor, and 0 otherwise. (SES_FLOOR_MATERIAL=5). It is considered as missing if SES_FLOOR_MATERIAL is missing.
    f. OTHER_FLOOR: a dummy variable equal to 1 if the household has a polished wood floor, and 0 otherwise. (SES_FLOOR_MATERIAL=6). It is considered as missing if SES_FLOOR_MATERIAL is missing.

Table 3 presents the results of regressing each variable above on the dummy variable ACT_SUBSIDY. Note that no household in the treatment or control groups has floor tiles, wood planks, polished wood floor, or other type of floor. Consequently the variables FLOOR_TILES, WOOD_PLANKS, POLISHED_WOOD, and OTHER_FLOOR are not included in Table 3.

The results do not suggest any significant difference between the treatment and the control group in terms of the type of floor.

Table 3 - Baseline Summary Statistics for Type of Floor (SES_FLOOR_MATERIAL)		
	Control group mean (1)	ACT_SUBSIDY

		(2)
EARTHEN_FLOOR	0.519	0.007
	(0.107)***	(0.060)
CEMENT_FLOOR	0.526	-0.007
	(0.093)***	(0.060)
Following Table 1 in Cohen et al (2015, page 623), the first column shows average values of characteristics for the control group. Column 2 shows regression coefficients and standard errors on the treatment group. All regressions include a full set of strata dummies. *** Significant at the 1 percent level. ** Significant at the 5 percent level.		

4. SES_WALL_MATERIAL. Description: Type of wall material. It includes the variables:
    a. STONE_WALL: a dummy variable equal to 1 if the household has stone walls, and 0 otherwise. (SES_WALL_MATERIAL=1). It is considered as missing if SES_WALL_MATERIAL is missing.
    b. BRICK_WALL: a dummy variable equal to 1 if the household has brick walls, and 0 otherwise. (SES_WALL_MATERIAL=2). It is considered as missing if SES_WALL_MATERIAL is missing.
    c. TIMBER_WALL: a dummy variable equal to 1 if the household has timber walls, and 0 otherwise. (SES_WALL_MATERIAL=3). It is considered as missing if SES_WALL_MATERIAL is missing.
    d. IRON_WALL: a dummy variable equal to 1 if the household has iron walls, and 0 otherwise. (SES_WALL_MATERIAL=4). It is considered as missing if SES_WALL_MATERIAL is missing.
    e. MUD_WALL: a dummy variable equal to 1 if the household has mud walls, and 0 otherwise. (SES_WALL_MATERIAL=5). It is considered as missing if SES_WALL_MATERIAL is missing.
    f. WOOD_WALL: a dummy variable equal to 1 if the household has wood walls, and 0 otherwise. (SES_WALL_MATERIAL=6). It is considered as missing if SES_WALL_MATERIAL is missing.
    g. CEMENT_WALL: a dummy variable equal to 1 if the household has cement walls, and 0 otherwise. (SES_WALL_MATERIAL=7). It is considered as missing if SES_WALL_MATERIAL is missing.
    h. OTHER_WALL: a dummy variable equal to 1 if the household has other types of walls, and 0 otherwise. (SES_WALL_MATERIAL=8). It is considered as missing if SES_WALL_MATERIAL is missing.

Table 4 presents the results of regressing each variable above on the dummy variable ACT_SUBSIDY. Note that no household in the treatment or control groups has timber walls, iron walls, wood walls, or other types of walls. Consequently the variables TIMBER_WALL, IRON_WALL, WOOD_WALL, and OTHER_WALL are not included in Table 4.

The results suggest that households in the treatment group are significantly more likely to have stone walls and cement walls, at a 5% significance level. Consequently, STONE_WALL and CEMENT_WALL are considered as part of the control variables in the replication.

Table 4- Baseline Summary Statistics for Type of Wall Material (SES_WALL_MATERIAL)

	Control group mean (1)	ACT_SUBSIDY (2)
STONE_WALL	0.526	0.353
	(0.093)***	(0.078)***
BRICK_WALL	0.529	-0.134
	(0.092)***	(0.083)
MUD_WALL	0.419	0.109
	(0.121)***	(0.081)
CEMENT_WALL	0.526	0.371
	(0.093)***	(0.085)***

Following Table 1 in Cohen et al (2015, page 623), the first column shows average values of characteristics for the control group. Column 2 shows regression coefficients and standard errors on the treatment group. All regressions include a full set of strata dummies.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.

5. SES_FUEL_TYPE. Description: Type of fuel. Note that no household in the treatment or control groups has information on this variable.
6. SES_NO_COWS. Description: Number of cows present in the household. The variable NUM_COWS is constructed using SES_NO_COWS, and is considered as missing if SES_NO_COWS is missing.

7. SES_NO_SHEEP. Description: Number of sheep present in the household. The variable NUM_SHEEP is constructed using SES_NO_SHEEP, and is considered as missing if SES_NO_SHEEP is missing.
8. SES_NO_GOATS. Description: Number of goats present in the household. The variable NUM_GOATS is constructed using SES_NO_GOATS, and is considered as missing if SES_NO_GOATS is missing.
9. SES_DHS_SCORE_POOLED. Description: Wealth Index (raw score), pooled.
10. SES_DHS_PERCENTILE_POOLED. Description: Wealth Index (quintile), pooled.

Table 5 presents the results of regressing each variable above on the dummy variable ACT_SUBSIDY. The results suggest that households in the treatment group are significantly more likely to have more sheep, at a 5% significance level. Consequently, NUM_SHEEP is considered as part of the control variables in the replication.

Table 5- Baseline Summary Statistics for NUM_COWS, NUM_SHEEP, NUM_GOATS, SES_DHS_SCORE_POOLED, and SES_DHS_PERCENTILE_POOLED

	Control group mean (1)	ACT_SUBSIDY (2)
NUM_COWS	0.525	0.001
	(0.094)***	(0.013)
NUM_SHEEP	0.512	0.038
	(0.091)***	(0.018)**
NUM_GOATS	0.518	0.018
	(0.093)***	(0.027)
SES_DHS_SCORE_POOLED	0.526	0.001
	(0.094)***	(0.026)
SES_DHS_PERCENTILE_POOLED	0.490	0.015
	(0.103)***	(0.018)

Following Table 1 in Cohen et al (2015, page 623), the first column shows average values of characteristics for the control group. Column 2 shows regression coefficients and standard errors on the treatment group. All regressions include a full set of strata dummies.
*** Significant at the 1 percent level.

** Significant at the 5 percent level.

According to the results in Tables 1 to 5, to avoid any confounding on the estimates, as control variables are included:

1. REFRIGERATOR
2. MOBILE
3. VIP_TOILET
4. COMPOSTING_TOILET
5. OTHER_TOILET
6. STONE_WALL
7. CEMENT_WALL
8. NUM_SHEEP

Note that Cohen et al (2015) use as a control variable the age of the household head. However, this variable is not available in the replication data, and consequently, it is not possible to analyze if households are different in this observable.

18. Indices

RR TEAM INSTRUCTIONS: *If any of the measured variables described in Section 17 will be combined into a composite measure (including simply a mean), describe in detail what measures you will use and how they will be combined. Please be sure this preregistration includes a link to a clearly commented script that constructs the index according to the narrative.*

Specific points to keep in mind (please also consult the Reviewer Criteria):

- *Does the preregistration specify each of the composite measures (e.g. mean scores, factor scores) that are needed for the focal analysis, and which of the measured variables in Section 17 are used in each one (e.g. the happiness, joy, and satisfaction items will be used to create the 'positive feelings' measure)?*
- *Does the preregistration link to a clearly commented script that constructs the indices according to the narrative description?*

No measured variables described in the "Measured variables" section need to be combined into a composite measure.

Analysis Plan

19. Statistical models

RR TEAM INSTRUCTIONS: *This section should describe in detail the analysis that will be performed to replicate the focal result. This analysis must align as closely as possible with the original study's analysis, even if you have identified limitations in the original study. The level of*

detail should allow anyone to reproduce your analyses from your description below. Examples of what should be specified: the model; each variable; adjustments made to the standard errors and to case weighting; additional analyses that are required to set up the focal analysis; and the software used.

Beyond the replication of the focal analysis from the original study, it is at your discretion to test the claim using other analytic approaches as a check of the robustness of the claim. The original test should be listed first and be clearly distinguished from any other tests. If you are testing additional confirmatory hypotheses, describe them in the same order as you numbered them in the "Hypotheses" section above and make clear reference to the specific hypothesis being tested for each.

Please provide a link to a clearly commented script that performs the analysis described in the narrative provided below. Our preference is that this be either an R script or a script from another language that similarly allows for open and reproducible analyses. Please let the SCORE team know if this is not possible. Please also test that the code runs without error on a random subset of 5% of the replication dataset, and provide verification that the code has produced a sensible result below by providing a screenshot of the output (please upload the screenshot to the OSF as well). Finally, please confirm that you have only developed and tested your analysis plan and code using 5% of the data.

Specific points to keep in mind (please also consult the Reviewer Criteria):

- *Does the preregistration specify which statistical model will be used to provide the 'focal evidence' for the SCORE test (e.g. a regression coefficient in a larger multiple regression model), and does it correspond closely to the model and evidence from the original study?*
- *Does the preregistration describe each variable that will be included in the focal analysis, and what role each variable has (e.g. dependent variable, independent variable)?*
- *Does the preregistration include a detailed specification of the focal analysis, including interactions, lagged terms, controls, etc., in both narrative form and in a clearly commented script?*
- *Does the preregistration verify that the code runs without error on a random subset of the replication dataset?*

This statement confirms that only 5% of the data have been randomly sampled in developing the analysis plan and code contained in this preregistration.

However, following Cohen et al (2015, pages 622-624), the control variables are selected based on characteristics that do not balance across the treatment and the control group. The idea is to avoid any confounding in the estimates due to a lack of balance across groups. Consequently, all available observations are used to identify any significant differences between the groups, as it is discussed in the "Measured variables" section

(Tables 1 to 5). Apart from that, only 5% of the data randomly sampled is used in developing the analysis plan and code contained in this preregistration.

Lines 258 to 260 of the Stata do-file "Cohen et al 2015 - Replication Analysis 5% random sample.do" generate a random sample that contains only 5% of the data for each state.

To replicate the analysis an Ordinary least squares (OLS) regression model is estimated. The dependent variable is a dummy equal to 1 if the subject took the ACT Artemether Lumefantrine for the current illness, and 0 otherwise. The focal independent variable is a dummy variable equal to 1 if the subject was given an ACT voucher subsidy, and 0 otherwise. A full set of strata fixed effects are included, as well as the following control variables:

1. REFRIGERATOR
2. MOBILE
3. VIP_TOILET
4. COMPOSTING_TOILET
5. OTHER_TOILET
6. STONE_WALL
7. CEMENT_WALL
8. NUM_SHEEP

Given the nature of the data, a survey weight is used to fit the OLS regression model. The survey weight is indicated by the variable WEIGHT. The estimation is restricted to individuals that get tested through the community health worker (CHW) program (i.e. MALTEST_WHERE=1) and for the 3 post-baseline surveys (i.e. WAVE!=0). As it is explained in the "Study Design" and "Sample size" sections, by design, the ACT subsidy is given to individuals who were tested through the CHW and got a positive result; and the focal dependent and independent variables are obtained from the 3 post-baseline surveys.

The software used is Stata 15.1 The analysis code is "Cohen et al 2015 - Replication Analysis 5% random sample.do", and the log file with the Stata session that demonstrates that this code works is "Cohen-et-al_Replication_5_Random_Sample.pdf".

20. Transformations

RR TEAM INSTRUCTIONS: *This section should describe how any of the measured variables or composite measures mentioned above will be transformed prior to the analyses listed in Section 19. These are adjustments made to variables **after** measurement or measure creation, and might include centering, logging, lagging, rescaling etc. Please provide enough detail such that anyone else could reproduce the transformations based on the description below. Please be sure this preregistration includes a link to a clearly commented script that performs the transformations described in the narrative provided below.*

Specific points to keep in mind (please also consult the Reviewer Criteria):

- *Does the preregistration specify which of the measured variables or composite measures will need to be transformed prior to the focal analysis?*
- *For each variable needing transformation, does the preregistration adequately describe the transformations, including any centering, logging, lagging, recoding, or implementation of a coding scheme for categorical variables?*
- *Does the preregistration link to a clearly commented script that performs each transformation?*

No measured variables or composite measures mentioned above will be transformed prior to the analyses in the "Statistical models" section.

21. Inference criteria

RR TEAM INSTRUCTIONS: *This section describes the precise criteria that will be used to assess whether the hypotheses listed above were confirmed by the analyses in Section 19. The default language below only applies to the test of the SCORE claim, **H***. It is at your discretion to describe the inferential criteria you will use for any additional analyses. They need not rely on p-values and/or the same alpha level we have specified for **H***.*

If the additional analyses will use multiple comparisons, the inference criteria is a question with few "wrong" answers. In other words, transparency is more important than any specific method of controlling the false discovery rate or false error rate. One may state an intention to report all tests conducted or one may conduct a specific correction procedure; either strategy is acceptable.

Criteria for a successful replication attempt for the SCORE project is a statistically significant effect (alpha = .05, two tailed) in the same pattern as the original study on the focal hypothesis test (**H***).

To test the SCORE claim, **H***, that ACT [artemisinin combination therapies] subsidies induce take-up of ACT, an Ordinary least squares (OLS) regression is estimated. Following Cohen et al (2015, page 626-627) the specification takes the form:

$$TOOK\_ACT_i = \delta + \gamma ACT\_SUBSIDY_i + \beta X_i + \lambda_{STRATA}$$

Where:

1. $TOOK\_ACT_i$: is a dummy variable equal to 1 if the subject i took the ACT Artemether Lumefantrine for the current illness, and 0 otherwise. It is the dependent variable.
2. $ACT\_SUBSIDY_i$: is a dummy variable equal to 1 if the subject i was given an ACT voucher subsidy, and 0 otherwise. It is the focal independent variable. The coefficient of interest is γ, which indicates the effect of an ACT [artemisinin combination therapies] subsidies on ACT take-up.

3. δ : constant term.
4. X_i : vector of control variables for subjects i's household. According to the "Measured variables" section, the covariates are:
    a. REFRIGERATOR
    b. MOBILE
    c. VIP_TOILET
    d. COMPOSTING_TOILET
    e. OTHER_TOILET
    f. STONE_WALL
    g. CEMENT_WALL
    h. NUM_SHEEP
5. β : vector of coefficients for X_i
6. λ_{STRATA} : strata fixed effects.

Note that to analyze the data as survey data, the survey weight is indicated by the variable WEIGHT. The estimation is restricted to individuals that get tested through the community health worker (CHW) program (i.e. MALTEST_WHERE=1) and for the 3 post-baseline surveys (i.e. WAVE!=0). As it is explained in the "Study Design" and "Sample size" sections, by study design, the ACT subsidy is given to individuals who were tested through the CHW and got a positive result; and the focal dependent and independent variables are obtained from the 3 post-baseline surveys.

Following Cohen et al (2015, page 627), robust standard errors clustered at the household are used.

22. Data exclusion

RR TEAM INSTRUCTIONS: *The section below should describe the rules you will follow to exclude collected cases from the analyses described in Section 19. Note that this refers to exclusions **after** the creation of the replication dataset; exclusion criteria that prevent a case from entering the replication dataset in the first place should be detailed in the 'Data Collection Procedure' section above. Please be as detailed as possible in describing the rules you will follow (e.g. What is the specific definition of outliers you will use? Exactly how many attention checks does a participant need to fail before their removal from the analytic sample?).*

Specific points to keep in mind (please also consult the Reviewer Criteria):

- *Does the preregistration comment on whether any cases included in the replication dataset will be excluded prior to data analysis?*
- *If yes, does the preregistration provided detailed instructions on how the exclusions will be performed (e.g. Is the definition of outlier provided? Is the number of attention checks failed before a participant is excluded specified?)*

The only reason to exclude observations in the analysis is due to missing values, as will be described in the “Missing data” section.

23. Missing data

RR TEAM INSTRUCTIONS: *The section below should describe how missing or incomplete data will be handled. Please be as detailed as possible in describing the exact procedures you will follow (e.g. last value carried forward; mean imputation) and any software required (e.g. We will use Amelia II in R to perform the imputation).*

Specific points to keep in mind (please also consult the Reviewer Criteria):

- *Does the preregistration comment on how missing or incomplete data will be addressed (e.g. casewise removal, missing data imputation)?*
- *If applicable, does the preregistration specify how many missing variables will lead to a case’s removal (e.g. If a subject does not complete any of the three indices of tastiness, that subject will not be included in the analysis.)?*
- *If applicable, does the preregistration describe how missing data imputation will be performed, including relevant software?*

For some respondents in the sample, some variables are reported as “Don’t remember” or “Missing.” (more details in the “Measured Variables” section). In those cases, the variable value is recorded as missing. Consequently, the respondent is excluded from the Ordinary Least Squares regression in the “Inference Criteria” section.

24. Exploratory analysis (Optional)

RR TEAM INSTRUCTIONS: *If you plan to explore your data set to look for unexpected differences or relationships, you may describe those tests here. An exploratory test is any test where a prediction is not made up front, or there are multiple possible tests that you are going to use. A statistically significant finding in an exploratory test is a great way to form a new confirmatory hypothesis, which could be registered at a later time. If any exploratory analyses involve additions to the data collection procedure beyond what was performed in the original study (e.g. additional items on the survey; running another condition in the experiment), please describe them below.*

25. Other

RR TEAM INSTRUCTIONS: *This section serves two purposes. First, please use this section to discuss any features of your replication plan that are not discussed elsewhere. Literature cited, disclosures of any related work such as replications or work that uses the same data, plans to make your data and materials public, or other context that will be helpful for future readers would be appropriate here. Second, please also re-surface any major deviations from earlier in the preregistration that you expect a reasonable reviewer could flag for concern. Give a*

summary of these deviations, focusing on larger changes and any possible challenges for comparing the results of the original and replication study.

Specific points to keep in mind (please also consult the Reviewer Criteria):

- *Does the preregistration reference other sections of the preregistration where substantial deviations from the original study have been described (including deviations due to differences in location or time compared to the original study)?*
- *Does the preregistration comment on plans to make the data and materials from the replication study public?*

Final review checklist

REVIEWER INSTRUCTIONS: *For the following questions, reviewers please indicate whether you can 'sign off' on the following items by adding a comment. You can update this response as the lab moves through revisions during the review period!*

- Included in this pre-registration are specific materials needed to create a replication dataset:
    - Is the final replication dataset that the research team constructed suitable for performing a high-quality, good-faith replication of the focal claim selected from the original study?
    - Is the procedure for constructing the final replication dataset sufficiently documented that an independent researcher could construct the same dataset following the procedures and code they lay out?
- Included with this pre-registration is a narrative description of how the replication dataset will be used to perform the focal replication analysis, as well as the specific analytic scripts/code/syntax that will be used:
    - Is the analysis plan (including code) that's documented in the preregistration consistent with a high-quality, good-faith replication of the focal claim selected from the original study?
    - Has the data analyst demonstrated that the analysis code works as expected on a random 5% of the final replication dataset?
- I have reviewed all sections of this pre-registration, and I believe it represents a good-faith replication attempt of the original focal claim.