Nazila Alinaghi

Risk sharing and transaction costs

A replication study of evidence from Kenya's mobile money revolution

February 2019

Replication Paper 22

Finance

International Initiative for Impact Evaluation

About 3ie

The International Initiative for Impact Evaluation (3ie) is an international grant-making NGO promoting evidence-informed development policies and programs. We are the global leader in funding, producing and synthesising high-quality evidence of what works, for whom, how, why and at what cost. We believe that using better and policy-relevant evidence helps to make development more effective and improve people's lives.

3ie Replication Paper Series

The 3ie Replication Paper Series is designed to be a publication and dissemination outlet for internal replication studies of development impact evaluations. Internal replication studies are those that reanalyse the data from an original paper in order to validate the results. The series seeks to publish replication studies with findings that reinforce or challenge the results of an original paper. To be eligible for submission, a replication study needs to be of a paper in 3ie's online Impact Evaluation Repository and needs to include a pure replication. 3ie invites formal replies from the original authors. These are published on the 3ie website together with the replication study.

The 3ie Replication Program also includes grant-making windows to fund replication studies of papers identified on our candidate studies list. Requests for proposals are issued one to two times a year. The candidate studies list includes published studies that are considered influential, innovative or counterintuitive. The list is periodically updated based on 3ie staff input and outside suggestions. The aim of the 3ie Replication Program is to improve the quality of evidence from development impact evaluations for use in policymaking and program design.

About this report

The Bill & Melinda Gates Foundation helped fund this report. All content, errors and omissions are the sole responsibility of the authors and do not represent the opinions of 3ie, its donors or its Board of Commissioners. Please direct any comments or queries to the corresponding author, Nazila Alinaghi, at nazila.alinaghi@vuw.ac.nz.

Suggested citation: Alinaghi, N, 2018. *Risk sharing and transaction costs: a replication study of evidence from Kenya's mobile money revolution*, 3ie Replication Paper 22. Washington, DC: International Initiative for Impact Evaluation (3ie). Available at: https://doi.org/10.23846/RPS0022

3ie Replication Paper Series executive editor: Marie Gaarder
Co-managing editor: Anna Heard
Production manager: Brigid Monaghan
Copy editor: Jaime L Jarvis
Proof reader: Yvette Charboneau
Cover design: John F McGill and Akarsh Gupta

Risk sharing and transaction costs: a replication study of evidence from Kenya's mobile money revolution

Nazila Alinaghi
Victoria University of Wellington

Replication Paper 22

February 2019

International Initiative for Impact Evaluation

Acknowledgements

First, I would like to thank the International Initiative for Impact Evaluation (3ie) and the Bill & Melinda Gates Foundation for providing me with the support to conduct this replication study. I wish to extent my great appreciation to 3ie's team for Replication Window 4 (Financial Services for the Poor) – Benjamin DK Wood, Eric Djimeu and Scott Neilitz – who helped me by providing invaluable guidance and support throughout the study. I also thank the internal and external reviewers and advisors and the selection panel for their constructive comments on the proposal and the final report. Finally, I am grateful to the original authors, professors William Jack and Tavneet Suri, who provided me with their code and data and were helpful in answering my questions.

Summary

M-PESA, a mobile phone–based technology for transferring money, provides a gateway to formal financial services for populations who otherwise would not have access to those services. In a 2014 study, William Jack and Tavneet Suri analyse a panel of 2,282 Kenyan households over the 2008–2010 period to estimate how M-PESA has enabled financial risk sharing. They focus on two types of negative shocks – an overall negative shock and an illness shock – and then analyse how family members and friends share financial resources during these adverse events. A key finding is that M-PESA users, relative to nonusers, are largely able to protect their consumption when faced with negative income shocks. They are more likely to receive remittances in the face of unexpected negative shocks, receiving both a greater number and higher value of remittances. They also receive remittances over greater distances and from wider networks.

This replication study uses pure replication, robustness checks and additional model specifications to re-examine this recent work on the impact of transaction costs on the ability of households to share risk using a mobile money innovation, M-PESA.

I conduct the twin strategies of push-button replication and pure replication. For the push-button replication, I use the code provided by the original authors. For the pure replication, I use an independent coding method. With only a few minor differences, push-button and pure replication results both support the original authors' findings. I also find that the results are robust to various consistency and robustness checks, such as propensity score matching and Tobit model specification.

Finally, I explore heterogeneous effects by comparing benefits across urban and rural residents. The latter are expected to particularly benefit from M-PESA, as they are more likely to be excluded from formal financial services. However, I cannot confirm the effect is more pronounced in urban versus rural segments.

Jack and Suri's original findings and my replication provide strong empirical evidence that M-PESA has had a positive impact on people's financial health. The financial benefits derived from market-based mobile money innovations can play a vital role in combating world poverty. Policymakers should take one crucial step: adopt enabling policies that allow competing firms to offer new technologies such as M-PESA.

Contents

List of figures and tables

Appendix tables

Abbreviations and acronyms

FE	Fixed effects
FSD Kenya	Financial Sector Deepening Kenya
HH	Household
IV	Instrumental variable
LM test	Lagrange Multiplier test
MFE-MC	Multilevel fixed effects, multiway clustering
OLS	Ordinary least squares
ROSCA	Rotating savings and credit association
SACCO	Savings and credit cooperative organization
SD	Standard deviation

1. Introduction

Mobile money, a financial inclusion instrument, has had a revolutionary impact on the lives of the poor, who previously had limited (if any) access to formal financial services. This recent innovation has expanded the reach of financial services beyond the realm of traditional ‘brick and mortar’ banks. Mobile money, as a means to improve access to financial services in low- and middle-income countries, is a key enabler to most of the United Nations Sustainable Development Goals,[1] which aim to end poverty, protect the planet and ensure prosperity for all (GSMA 2017).

A recent study by Suri and Jack (2016) found that within eight years, M-PESA lifted 194,000 households – 2 per cent of the Kenyan population – out of poverty, helping people develop greater financial resilience and savings, especially among female-headed households. M-PESA also enabled 185,000 women to move out of subsistence farming into business or retail occupations. Thus, it contributes towards Sustainable Development Goal 1, eradicating poverty, and Sustainable Development Goal 5, achieving gender equality and economic empowerment of women (UNCDF 2018).

Although access to financial services has clearly grown in developing nations, evidence of its potential economic effects has been less clear. As a result, a growing body of recent academic research has focused on understanding the economic impact of mobile money as a poverty alleviation tool in developing economies (Aker and Mbiti 2010; Mbiti and Weil 2011; Jack and Suri 2011; Suri and Jack 2016; Munyegera and Matsumoto 2016).

In low- and middle-income countries, high income risk is an inevitable part of life (Alderman and Paxon 1992). In addition to climate risks and economic fluctuations, a large number of idiosyncratic shocks make these households vulnerable to severe hardship. In Kenya, a large number of households are exposed to a variety of risks, including crop failure (as a result of drought, flood and other climate events), illness or job loss. In order to control risk and cope with these income shocks, households regularly make use of a variety of strategies. Given poor access to formal insurance, informal networks, such as family members and friends, can play a crucial role in sharing risks in the event of any negative shocks (De Weerdt and Dercon 2006; Fafchamps and Lund 2003).

M-PESA,[2] Kenya’s world-leading mobile phone–based service, has facilitated a range of financial transactions, including deposit, withdrawal and transferring capabilities. Despite the wide range of services offered, person-to-person transfers dominate M-PESA use, mainly due to the growing pace of the rural-to-urban migration, which is accelerating demand for secure, fast and affordable means of sending remittances home (Aker and Mbiti 2010).

Jack and Suri (2014) explore the impact of mobile money on informal risk sharing. They use panel data to analyse how mobile money can improve the ability of households to

---

[1] The Sustainable Development Goals are a collection of 17 global goals to transform our world by the United Nations.
[2] M-PESA: M stands for mobile and *pesa* is Swahili for money.

smooth consumption in response to negative idiosyncratic shocks. They find that the introduction of M-PESA facilitates the redistribution of finances across geographical distances. The success of M-PESA in improving households' ability to spread risk is mainly attributed to a reduction in transaction costs. As a result, M-PESA has helped lift Kenyan households out of poverty (Suri and Jack 2016).

To better understand the robustness of this promising evidence, this paper aims to investigate key findings Jack and Suri report in their influential *American Economic Review* article, 'Risk Sharing and Transaction Costs: Evidence from Mobile Money Revolution in Kenya' (2014). My replication proceeds as follows. In Section 2, I explain the twin replication strategies of push-button replication and pure replication. In this section, I explain the data, methods and assumptions I use to conduct the replication research. In Section 3, I explore some of the alternate strategies to determine the robustness of the results to the analysis method. In Section 4, by changing the focus of analysis from the entire population (including rural and urban residents) to the rural segment of the population supposedly excluded by formal financial services, I am able to assess whether M-PESA has abandoned its initial promise, 'banking for the unbanked'. I then provide some limitations and conclusions at the end of the report.

I have three main findings. First, with only a few minor differences, both the push-button and pure replication results support the original authors' findings. Second, I find that their results are robust to various consistency tests and robustness checks, such as propensity score matching and Tobit model specification. Third, I am not able to confirm that the effect of mobile money is more pronounced for urban than rural residents. My replication provides strong empirical evidence confirming that M-PESA has had a positive impact on people's financial health.

2. Replication

2.1 Push-button replication

My replication study (Alinaghi and Reed 2017) starts with push-button replication, in which one attempts to 'push a button' and reproduce the published tables in the original study using the code provided by the authors. I downloaded the original data set and code available on the paper's webpage at the *American Economic Review*.[3] Although the original paper used two rounds of household survey data, the extract available on the webpage did not include data from the first round. To obtain access to these data, I contacted the Financial Sector Deepening Trust Kenya (FSD Kenya), as instructed by the authors' webpage. Once I received permission from FSD Kenya, Jack and Suri provided the final extract used in the original paper.

With only a few minor differences, my push-button replication was able to exactly reproduce the authors' original results. A few tables (such as part of Table 2, Table 7A and two tables in the Web Appendix) could not be reproduced, since I was unable to obtain the associated data and code from the authors.

[3] Data and code can be found at https://www.aeaweb.org/articles?id=10.1257/aer.104.1.183.

2.2 Pure replication

Although a push-button replication seeks to duplicate a paper's findings using the original data and code provided by the authors, pure replication attempts to reproduce the paper's findings by reconstructing data and code, following the descriptions and explanations supplied in the text.

Two common methodologies for pure replication are code auditing and independent coding. I chose to implement the latter. To do so, I read the description associated with each table and attempted to write the code to reproduce the corresponding findings. In some cases, the description provided in the paper gave insufficient detail. When this occurred, I referred to the original code to better understand the authors' intent.

2.2.1 The data

Jack and Suri's study is based on a large panel household survey, conducted over a three-year period. The initial survey was conducted on 3,000 households from August–October 2008. Two follow-up surveys were conducted from October 2009–January 2010 and May–August 2010. Attrition was high. Although the original study consisted of 3,000 households, the number of households interviewed for the second and third rounds fell to 2,105 and 1,531. Notably, 265 households missing from the second round reappeared in the third round.

The high attrition rate created a challenge, as only 1,266 households appeared in all three rounds of the survey. As a result, the authors decided to compress the panel into two rounds, taking the households included in the first and second rounds and adding to them the 265 households that appeared in the first and third rounds, but not the second round. This allowed a balanced, two-period panel with a potential maximum of 2,370 households, or 4,740 observations.

The actual extract provided by the original authors consisted of 5,281 observations, with 2,999 observations from the first survey, 2,017 from the second and 265 from the third. The last two rounds were combined and households not available in the second period were dropped from the first round, resulting in a balanced, two-period panel extract of 2,282 households.

2.2.2 Brief description of observational study

Jack and Suri motivate their study using a theoretical framework to demonstrate how transaction costs can play a significant role in risk sharing. To empirically study how M-PESA affected risk sharing, they use a difference-in-difference analysis to compare changes in consumption between M-PESA users and nonusers in response to negative income shocks. They estimate the following specification:

$$C_{ijt} = \alpha_i + \gamma Shock_{ijt} + \mu User_{ijt} + \beta User_{ijt} \times Shock_{ijt} + \theta X_{ijt} + \eta_{jt} + \pi_{rt} + \varepsilon_{ijt}, \quad (1)$$

where C_{ijt} is annual per capita consumption for household i in location j in period t, α_i is a household fixed effect, $Shock_{ijt}$ is a dummy variable equal to 1 if the household reported experiencing a negative shock, $User_{ijt}$ is a dummy taking the value of 1 if the household was an M-PESA user, X_{ijt} is a vector of controls and η_{jt} and π_{rt} are location-by-time and rural-by-time dummy variables.

Since remittances is the hypothesised channel by which households can smooth consumption, Jack and Suri estimate another version of the above estimation:

$$r_{ijt} = \alpha_i + \gamma Shock_{ijt} + \mu User_{ijt} + \beta User_{ijt} \times Shock_{ijt} + \theta X_{ijt} + \eta_{jt} + \pi_{rt} + \varepsilon_{ijt} \ , \quad (2)$$

where r_{ijt} is a measure of remittances over the previous six months.

To further control for any observable characteristics that could affect the ability of households to smooth consumption, Jack and Suri provide a specification in which the interaction of the negative shocks with all observable covariates have been included:

$$C_{ijt} = \alpha_i + \gamma Shock_{ijt} + \mu User_{ijt} + \beta User_{ijt} \times Shock_{ijt} + \theta^S X_{ijt} \times Shock_{ijt} + \theta^M X_{ijt} + \eta_{jt} + \pi_{rt} + \varepsilon_{ijt} \quad (3)$$

As a further robustness check, Jack and Suri use agent rollout data to estimate the following specification:

$$C_{ijt} = \alpha_i + \gamma Shock_{ijt} + \vartheta Agent_{ijt} + \beta Agent_{ijt} \times Shock_{ijt} + \theta^M X_{ijt} + \eta_{jt} + \pi_{rt} + \varepsilon_{ijt}, \quad (4)$$

where $Agent_{ijt}$ is a measure of access to an M-PESA agent.

2.2.3 Replicated results

I re-estimated the results in Jack and Suri's Tables 1–9. Additionally, I reproduced their Tables 1–3, which appear in their Online Appendix. A further two tables were published in a web appendix; however, I was unable to reproduce these due to a lack of data. My replication results are comparable with the original findings. Technically, however, my results must be categorized as incomplete, as I was unable to replicate all the tables due to missing data. Each of the reproduced outcomes is discussed below.

Table 1A–1C: summary statistics

I begin my pure replication by reproducing summary statistics for the full sample reported in the first table of the original paper. Appendix Table A1[4] compares variables over the first two rounds of the survey (2008 and 2009). My findings are identical to those reported in Table 1A in the original paper. However, a notable omission from Table 1 reported in the original paper is the age of the household head. If wealth correlates with the age of the household head, which is plausible, more strategic remittances would be expected.

Table 1B in the original study provides additional information on remittances over the two survey rounds.[5] Appendix Table A2 reports the results obtained when I use my code, following Jack and Suri's descriptions in the text. The grey-highlighted section indicates a difference in the values produced by my code and the values reported in their paper.

[4] The results corresponding to pure replication are reported in Appendix Table A1–Table A25.

[5] Although Jack and Suri provide some detail on the nature of remittances in the Table 1B, further information, such as the recipients of remittances (to whom the remittances were sent), would have been helpful. This information is available in the raw data sets. The problem is taking these data and matching them to the extracts used for the *American Economic Review* analysis. It was not possible to do this without further assistance from the original authors. Other information, such as the timing of remittances relative to the timing of the shocks, was not available in the raw data.

Upon inspection of the code, I realized the discrepancy arose because the authors' code converted zeroes to missing values for the first five variables. Although I am not sure whether they cleaned the data at this stage, once I adjust my code accordingly, I reproduce Table 1B exactly (Appendix Table A3).

Table 1C in the original study is based on an extract taken from the first survey round. This table compares transaction costs for various types of transfers. I was able to exactly replicate this table (Appendix Table A4). I note that other than hand delivery and bus delivery by a friend or relative, the lowest transaction costs are associated with M-PESA.

Table 2: agents' characteristics
In addition to surveying households, Jack and Suri also report information about access to M-PESA agents. These results are reported in Table 2 of the original paper. The associated replication is reported in the Appendix Table A5. I was able to exactly replicate Panels A and C of Table 2 in the original paper. However, I was unable to reproduce Panels B (agent distribution), because I could not obtain these data from the authors.[6]

Table 3: correlation of shock measures
Table 3 in the original paper summarizes correlations between the two shock measures (overall and illness shocks) and a number of household-level variables. Appendix Table A6 reports my respective replication. Note that 'Agents within 1, 2 and 5 km' actually refers to the square root of the number of agents. The reason for using the square root is that the number of agents has a long right tail. The grey-highlighted cells indicate minor differences in the partial R^2 produced by Jack and Suri's code and the values reported in their table.

Basic specification
Tables 4A and 4B: basic difference-in-difference results
Jack and Suri's main results are reported in Table 4 of their paper, which is based on estimating equation (1) above. The main conclusion to draw from this table is that M-PESA users are better able to smooth their consumption in the presence of negative shocks (the coefficient on 'User × negative shock'). Appendix Table A7 presents my replication results. These matched exactly after I adjusted for a discrepancy between the authors' code and their description of the table.

In the process of replicating Table 4, I uncovered a discrepancy between their code and their description of Table 4. In the notes below the table, Jack and Suri state, 'Heteroscedasticity - robust standard errors in brackets'. However, the standard errors reported in the first column ('OLS') are conventional ordinary least squares (OLS) standard errors. When I estimated robust standard errors consistent with the table notes, the variable 'User × negative shock' lost its statistical significance. The heteroskedasticity-robust standard error was 0.0717, with a p-value of 0.201.

Table 4B in the original paper checks the results of Table 4A for robustness, varying the sample dependent variable and control variables. I was able to exactly match Jack and

[6] There is a minor difference for the total number of agents (light grey–highlighted area). The dark grey–highlighted cells show the cases I was unable to replicate due to data unavailability.

Suri's results (Appendix Table A8). According to the title of column 4 in Table 4B and the corresponding explanation on page 20 of the original paper, one would be led to believe that the only district excluded from the sample for this column was Mombasa, the second largest city in Kenya. However, it turns out that the sample corresponding to column 4 excludes both Mombasa and Nairobi. As a result, I re-estimated the specification of column 4, excluding Mombasa but keeping Nairobi. The results are reported in column '4-new' (the grey-highlighted column). The coefficient on 'User × negative shock' is consistent with the other robustness checks of Table 4B.

Table 4C: results for health shocks
Table 4C in the original paper continues to check for robustness by exploring the impact of health shocks. My replication results, based on the authors' descriptions, are given in Appendix Table A9. Although my results are generally consistent with Table 4C, I obtained different estimates. No significant difference can be observed between the results reported in Appendix Table A9 and Table A10. Whereas the whole sample is considered in the former table, Nairobi is excluded from the latter. Closer investigation uncovered that the results in the source table excluded Nairobi. However, this was not mentioned in the text. When I excluded Nairobi, I obtained results identical to Table 4C in the original paper (Appendix Table A10).

Mechanisms

Table 5A: mechanisms
One of the main channels by which households share risks is through remittances. Whereas equation (1) focuses on consumption outcomes, equation (2) tests whether remittances respond to negative shocks in the manner predicted by the theory. Table 5A in the original paper reports the results from estimating equation (2), and my replication results are reported in Appendix Table A11.

The results reported in the first two columns show the likelihood of receiving remittances. The third column represents the frequency of remittances. In column 4, the dependent variable is total remittances received by households. To account for the extreme positive skewness of total remittances, Jack and Suri worked with its square root. In addition, they excluded outliers from their sample.[7] Results from excluding Mombasa from the sample are reported in columns 5 and 6. The last two columns focus on illness shocks. I was able to exactly reproduce their results.

The estimates reported in Appendix Table A11 confirm that M-PESA users are more likely to receive remittances in the event of a negative shock. All of the 'User × shock' coefficients have the expected (positive) sign and all are significant to at least the 10 per cent significance level.

Appendix Table A12 in the original paper reports estimates of the impact of M-PESA on the size and number of networks households have access to in the event of negative shocks. Three measures of remittance networks are used: (i) the average distance remittances travelled to reach the target household, (ii) the number of different

[7] According to Jack and Suri's code, outliers are defined as those who receive more than 42,000 Kenyan shillings.

relatives/friends from whom remittances were received and (iii) the fraction of networks remitting. Again, I was able to exactly replicate the original results.

Results using agent data

To control for potential endogeneity associated with being an M-PESA user, Jack and Suri used agent rollout data, replacing user with agent in the estimated specifications (see equation (4)). The corresponding results are reported in their Tables 6A and 6B. My replication results are reported in Appendix Table A13 and Table A14.

The key variable in these tables is the agent variable. It generally measures the number of agents within a certain distance of the household (1 kilometer, 2 kilometer, 5 kilometer and 20 kilometer). In columns 5 and 6 of Appendix Table A14, the agent variable is measured by the log of distance of the closest agent, where distance is measured in meters. The 'Agents × shock' coefficient always has the expected sign and is generally significant. I was able to exactly reproduce the original results in both tables.

Appendix Table A15 reports the correlation between agent rollout and household-level observables. The exogeneity of the agent variable is supported by the general lack of significant correlations between the presence of nearby agents and household variables. The grey-highlighted cells indicate cases in which the replicated results differ from the original estimates. The discrepancies do not affect any of the qualitative conclusions from this table.

Falsification test

To further examine whether there is a systematic relationship between the rollout of M-PESA agents and household characteristics, Jack and Suri performed a falsification test using data prior to the introduction of M-PESA (1997–2007). As they explain in the paper, they make use of an entirely rural survey, in which the total consumption data had not been collected. As a result, they change the focus of analysis from total consumption to a limited number of items in consumption, including maize consumption and some other components of food consumption. The results are reported in Table 7A in the original paper. Unfortunately, I was unable to obtain these data from the authors and, therefore, was unable to confirm their estimates (Appendix Table A16).[8]

Table 7B in the original paper tailors the post-M-PESA sample to match the pre-M-PESA sample used in the falsification test of Table 7A. My replication of these results are reported in Appendix Table A17. I was able to exactly reproduce these estimates.

Attrition

A major concern in Jack and Suri's study is that their results may be affected by the high attrition rate in the panel surveys. This rate was initially about 33 per cent. As explained above, the interview strategy in the follow-up surveys was designed to reduce this rate to

---

[8] With respect to Table 7, Professor Suri communicated in private correspondence that the data were proprietary and she was, therefore, unable to share them. However, there was no indication of this in Jack and Suri's published work. In fact, the accompanying 'Data Read Me' statement states that the Tegemeo Institute were in the process of making their data publicly available. Unfortunately, despite numerous attempts to contact the institute, I was unable to elicit a response.

a subsequent 24 per cent. For the non-Nairobi sample, the attrition rate was 18 per cent. Although the authors provide a discussion of the comparability of their attrition rates with other studies (footnote 19, page 192), they also conduct additional robustness checks to show that their results were not qualitatively affected by attrition.

The first set of results are reported in their Table 8A. The results in the table come from a multivariate regression, in which the dependent variable is a binary variable indicating whether the household did not attrite from the sample. My replication results are reported in Appendix Table A18. I am able to replicate all their results, with two exceptions: the estimates for 'HH has a ROSCA account' and 'HH has a SACCO account' are mistakenly switched in the results reported by Jack and Suri (see grey-highlighted cells).

The original authors further investigate the effects of attrition in their Table 8B. My corresponding replication results are reported in Appendix Table A19. The first three columns address attrition, using the weighting strategy developed by Fitzgerald and colleagues (1998). The remaining columns restrict the sample to communities with low attrition rates (the average attrition rate here is about 7 per cent). Again, I am able to exactly reproduce their estimates.

Instrumental variable results

The authors take a further look at endogeneity in their Table 9. Here they address potential endogeneity concerns with M-PESA use by employing an instrumental variable (IV) procedure. Specifically, they instrument the two variables: M-PESA use and the M-PESA interaction term. My replication results are reported in Appendix Table A20. I was able to exactly reproduce their estimates.

Online appendix

To further support the findings in their paper, Jack and Suri provided additional analyses in their Online Appendix Tables 1–3. Appendix Table A21 reproduces their Online Appendix Table 1, which provides additional summary statistics by adoption status. They categorize three groups of households: early adopters, late adopters and non-adopters. Of the three, early adopters – those who had adopted M-PESA in period 1 – are wealthier, more educated and more likely to use formal financial services.

Appendix Table A22 replicates Online Appendix Table 2 from the original paper. It reports the estimated effects of negative shocks on consumption and risk sharing after controlling for remittances. I was able to exactly reproduce their estimates. Note that controlling for remittances reduces the magnitude of the 'User × Shock' coefficient and renders the estimates insignificant. Appendix Table A23 reproduces Online Appendix Table 3 from the original paper. The data come from rounds 3 and 4 of the survey. My replication exactly reproduces Jack and Suri's estimates.

Web appendix

The original paper mentions that some additional results using the unbalanced, three-period panel data are posted online.[9] Unfortunately, I was unable to obtain these data

[9] These results are available at:
http://mitsloan.mit.edu/shared/ods/documents/?DocumentID=2481

from the authors. Thus, I was not able to replicate the associated two tables from their Web Appendix.

2.2.4 Push-button and pure replication conclusions

My analysis employed a push-button replication and the ‘independent coding’ method of pure replication to study Jack and Suri (2014). I was able to reproduce all of the original paper’s main findings, although I did discover some minor differences in the code and tables.

Although the authors provided us with most of the data I needed to reproduce their results, they did not provide all the data required for a complete replication. As a result, I was unable to replicate all of their tables. Accordingly, my pure and push-button replication should be categorized as comparable but incomplete.

Jack and Suri’s original findings and my replication provide strong empirical evidence that M-PESA has had a positive impact on people’s financial health. The financial benefits derived from such mobile money innovations can play a critical role in combatting world poverty. To provide useful insights for future policy-making, there is a need for further robustness and sensitivity analyses of the original findings, as discussed in the following measurement and estimation analysis and theory of change analysis sections.

3. Measurement and estimation analysis

Although Jack and Suri (2014) conducted a thorough analysis, further robustness checks can be conducted to investigate the sensitivity of their findings beyond the checks originally performed. In this section, I first look at the data and check the existence of outliers. I then assess the strength of the original results to alternative statistical and estimation methods.

3.1 Detecting outliers

An outlier is considered as an observation that lies far outside the norm of a variable. The deleterious effects of outliers on statistical analysis have been discussed extensively by Osborn and Overbay (2004). The presence of outliers increases error variance and reduces the power of statistical tests. They can also distort regression estimates. There are many reasons why outliers may arise, but a common cause is human error. Survey data seem particularly vulnerable to human error. For these reasons, an analysis of outliers may prove insightful. My analysis employs a number of procedures available in Stata for the detection and replacement or removal of outliers: discrepancy measures, leverage measures and influence measures (Williams 2016). I consider both deletion of outliers and winsorising of extreme values.

3.1.1 Pre-estimation procedures

Missing values

Missing values are generally excluded from the regression analysis. However, since household surveys conducted between late 2008 and early 2010 (three rounds or two periods), two observations per household are available in the data set. As a result, the households with missing values in one round are not fully dropped from the sample. This could confound the results, as the change in the response of consumption and/or remittances over this period is of interest.

To address this concern, households with missing values in one round are dropped from the sample.[10] Table 1 summarizes the results of this robustness check for OLS specification. It is worth mentioning that all the panel specifications are robust to this alteration. As Table 1 shows, the replication results are consistent with the original findings.

Table 1: Basic difference-in-differences results, removing missing values

Total consumption (Full sample)		
	Original OLS (1)	Replication OLS (2)
M-PESA user	0.5730***	0.5725***
	[0.0377]	[0.0377]
Negative shock	–0.2111***	–0.2120***
	[0.0381]	[0.0381]
User × negative shock	0.0917*	0.0929*
	[0.0506]	[0.0507]
Demographic controls		
Controls + interactions		
Time FE	Yes	Yes
Time × location FE		
Observations	4,562	4,560
Negative shock	–0.1593***	–0.1594***
	[0.0252]	[0.0252]
Shock, users	–0.1194***	–0.1191***
	[0.0335]	[0.0335]
Shock, nonusers	–0.2111***	–0.2120***
	[0.0381]	[0.0381]
Shock, nonusers | user Xs		
Mean of user	0.5656	0.5658

Notes: Dependent variable is log total household consumption per capita. Heteroscedasticity-robust standard errors are in brackets. ***, **, * Significant at 1, 5 and 10 per cent levels.

Extreme values

Extreme values are defined as observations in a sample that are highly unlikely to occur. A user-written command in Stata (Nick Cox's extremes command, Cox 2004) provides an easy way of identifying the cases with the most extreme values (upper and lower bounds). Table 2 reports the lowest and highest values, respectively. To decide how to handle these extreme values properly, I check the survey data for each round provided by the original authors.[11]

[10] There are two households with missing values in one round. I dropped these households (four observations) from the sample.

[11] These data sets are available at: https://dataverse.harvard.edu/dataverse.xhtml?alias=mobilemoney

A closer look at the household with ID 1428005 reveals that the only reported expenditure for this household is monthly expenditures. Neither weekly expenditures nor annual expenditures are reported for this household; as a result, the total expenditure is relatively low (6.174) with respect to its mean (10.859). The weekly expenditures, according to the codebook, are defined as expenditures on foods.[12] Therefore, I drop households for which no weekly expenditures are recorded.[13]

As for other extreme values, except two cases highlighted in the bottom panel of Table 2, all else seems correct. The potential reasons for reporting as an unusual case are stated in the last column. The extreme values can be easily detected by looking at the distribution demonstrated in Figure 1. They are mainly located below the value of about 8 and above 14 (the response variable is per capita consumption).

Table 2: Extreme values

Lowest (highest) values			
Household ID	Log total household consumption per capita	Rounds	Potential reasons
Panel A. Households with lowest values			
1428005	6.173786	2	No weekly and annual expenditures
910031	7.361799	1	No annual expenditures, confirmed
910005	7.74327	2	Household size (10 members), confirmed
910048	7.821643	1	No annual expenditures, confirmed
1650072	7.855545	3	Confirmed
Panel B. Households with highest values			
1334029	14.42262	1	Large monthly expenditures 1,196,400 Couldn't confirm (mine: 14.44525)
1436028	14.44179	1	Large monthly expenditures 1,085,700, confirmed
1636031	14.44186	2	Large monthly expenditures 1,193,760 Couldn't confirm (mine: 14.50654)
1337014	14.50009	1	Large annual expenditures 1,031,000, confirmed
1337048	14.81708	1	Large annual expenditures 2,070,000, confirmed

[12] Total expenditures are defined as the sum of weekly, monthly and yearly expenditures.
[13] It might be acceptable not to have monthly and/or yearly expenditures, but it seems unusual not to spend on foods. Households with zero or missing weekly expenditures are as follows: 1428005, 1433001 and 1280060 for zero values and 1603020, 1295003 and 112118 for missing values.

Figure 1: Distribution of per capita consumption, sample size 4,564

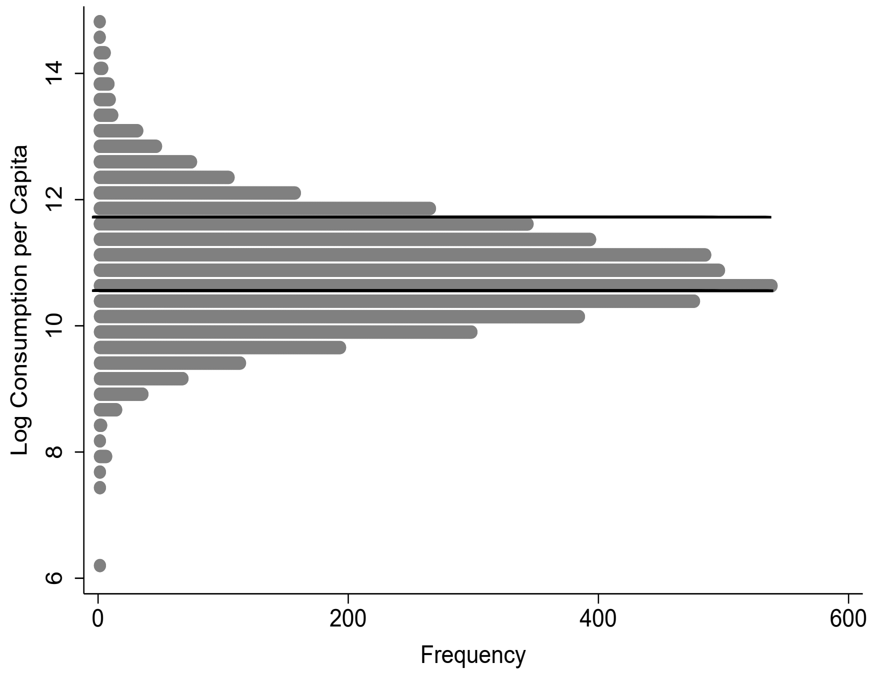


Table 3 presents the results of Jack and Suri's basic specifications, once I drop all the households with zero or missing weekly expenditures. The signs and sizes of the estimates are consistent with those reported in Table 4A.

Table 3: Basic difference-in-differences results, removing missing values for weekly expenditures

	Total consumption (Full sample)									
	OLS (1)		Panel (2)		Panel (3)		Panel (4)		Panel (5)	
	Original (1)	Replication (2)	Original (3)	Replication (4)	Original (5)	Replication (6)	Original (7)	Replication (8)	Original (9)	Replication (10)
M-PESA user	0.5730*** [0.0377]	0.5722*** [0.0377]	0.0520 [0.0481]	0.0501 [0.0481]	0.0456 [0.0469]	0.0450 [0.0469]	–0.0223 [0.0484]	–0.0231 [0.0485]	–0.0088 [0.0449]	–0.0106 [0.0448]
Negative shock	–0.2111*** [0.0381]	–0.2122*** [0.0381]	–0.0668 [0.0491]	–0.0653 [0.0491]	–0.0727 [0.0468]	–0.0711 [0.0468]	0.2872 [0.1762]	0.2901 [0.1760]	0.2673 [0.1799]	0.2777 [0.1797]
User × negative shock	0.0917* [0.0506]	0.0927* [0.0506]	0.1093* [0.0616]	0.1088* [0.0616]	0.1320** [0.0594]	0.1307** [0.0595]	0.1749*** [0.0663]	0.1736*** [0.0664]	0.1483** [0.0599]	0.1477** [0.0599]
Demographic controls					Yes	Yes	Yes	Yes	Yes	Yes
Controls + interactions							Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time × location FE			Yes	Yes	Yes	Yes			Yes	Yes
Observations	4,562	4,552	4,562	4,552	4,562	4,552	4,545	4,535	4,545	4,535
Negative shock	–0.1593*** [0.0252]	–0.1597*** [0.0252]	–0.0050 [0.0305]	–0.0037 [0.0304]	0.0019 [0.0292]	0.0029 [0.0293]	0.0022 [0.0286]	0.0018 [0.0286]	–0.0059 [0.0280]	–0.0045 [0.0280]
Shock, users	–0.1194*** [0.0335]	–0.1195*** [0.0334]	0.0425 [0.0379]	0.0435 [0.0379]	0.0592 [0.0370]	0.0596 [0.0370]	0.0518 [0.0383]	0.0508 [0.0383]	0.0460 [0.0355]	0.0468 [0.0355]
Shock, nonusers	–0.2111*** [0.0381]	–0.2122*** [0.0381]	–0.0668 [0.0491]	–0.0653 [0.0491]	–0.0727 [0.0468]	–0.0711 [0.0468]	–0.0626 [0.0447]	–0.0623 [0.0447]	–0.0737* [0.0429]	–0.0716* [0.0429]
Shock, nonusers | user Xs							–0.1230** [0.0549]	–0.1229** [0.0549]	–0.1024** [0.0502]	–0.1009** [0.0502]
Mean of user	0.5656	0.5662	0.5656	0.5662	0.5656	0.5662	0.5661	0.5667	0.5661	0.5667

Notes: Dependent variable is log total household consumption per capita. Heteroscedasticity-robust standard errors are in brackets. Controls: household demographics; household head education and occupation; use of bank accounts, SACCOs and ROSCAs; cell phone ownership. Interactions refer to interactions of the controls with the shock. When interactions are included, the overall effect of shock is evaluated at the mean of the covariates. All of these specifications controls for time fixed effects (Time FE). ***, **, * Significant at 1, 5 and 10 per cent levels.

3.1.2 Post-estimation procedures

Several post-estimation checks can help us to identify outliers. When I compare the ‘top 10’ post-estimation outliers with the pre-estimation outliers, 7 out of 10 appear on both lists (the grey-highlighted cells).

Table 4: Post-estimation outliers

Observation number	Household ID	Standardized residuals	R-student residuals	Leverage measure	Cooks distance measure	DFBETA measure
1615	1428005	–5.111664	–5.125822	0.001641	0.0071581	0.1513229
1628	910005	–3.799484	–3.8051	0.0008888	0.0021403	0.0060666
1639	910031	–3.763189	–3.768638	0.001199	0.0028334	–0.0068932
1622	910048	–3.215631	–3.218933	0.001199	0.0020689	–0.0058877
310	1123043	–3.211497	–3.214785	0.0050664	0.0087532	0.1067277
2511	1448033	4.046562	4.053409	0.001199	0.0032762	0.0074141
185	1636031	4.175604	4.183158	0.0008888	0.002585	–0.0066693
3119	1337048	4.279891	4.288051	0.0010445	0.0031921	0.0845722
3260	1365022	4.531851	4.541602	0.001641	0.0056263	–0.1340757
2419	1436028	4.667282	4.677966	0.001199	0.0043584	0.0085564

In Table 5, I show results for the sample, excluding the common outliers explained earlier. The coefficient on the interaction of interest in column 4 – which shows the panel specification, including the control for rural by time dummies and location by time dummies – is no longer significant at the 10 per cent level. The main difference between this column (column 4) and the following column (column 6) is demographic controls. Although the former is not controlling for household demographics, the latter controls for it. Thus, it could indicate that when I exclude the extreme values, the results are quite sensitive to demographic controls. The replication results for the other columns are consistent with the original findings.

Table 5: Basic difference-in-differences results, removing common pre-estimation and post-estimation outliers

	Total consumption – full sample									
	OLS (1)		Panel (2)		Panel (3)		Panel (4)		Panel (5)	
	Original (1)	Replication (2)	Original (3)	Replication (4)	Original (5)	Replication (6)	Original (7)	Replication (8)	Original (9)	Replication (10)
M-PESA user	0.5730***	0.5586***	0.0520	0.0606	0.0456	0.0551	–0.0223	–0.0213	–0.0088	–0.0088
	[0.0377]	[0.0370]	[0.0481]	[0.0474]	[0.0469]	[0.0462]	[0.0484]	[0.0484]	[0.0449]	[0.0444]
Negative shock	–0.2111***	–0.2109***	–0.0668	–0.0436	–0.0727	–0.0502	0.2872	0.3350	0.2673	0.3195
	[0.0381]	[0.0374]	[0.0491]	[0.0475]	[0.0468]	[0.0452]	[0.1762]	[0.1725]	[0.1799]	[0.1774]
User × negative shock	0.0917*	0.0939*	0.1093*	0.0920	0.1320**	0.1143**	0.1749***	0.1749***	0.1483**	0.1515**
	[0.0506]	[0.0497]	[0.0616]	[0.0604]	[0.0594]	[0.0583]	[0.0663]	[0.0661]	[0.0599]	[0.0594]
Demographic controls					Yes	Yes	Yes	Yes	Yes	Yes
Controls + interactions							Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time × location FE			Yes	Yes	Yes	Yes			Yes	Yes
Observations	4,562	4,548	4,562	4,548	4,562	4,548	4,545	4,531	4,545	4,531
Negative shock	–0.1593***	–0.1577***	–0.0050	0.0084	0.0019	0.0146	0.0022	0.0121	–0.0059	0.0075
	[0.0252]	[0.0247]	[0.0305]	[0.0297]	[0.0292]	[0.0285]	[0.0286]	[0.0280]	[0.0280]	[0.0273]
Shock, users	–0.1194***	–0.1169***	0.0425	0.0483	0.0592	0.0642	0.0518	0.0530	0.0460	0.0506
	[0.0335]	[0.0328]	[0.0379]	[0.0376]	[0.0370]	[0.0367]	[0.0383]	[0.0381]	[0.0355]	[0.0352]
Shock, nonusers	–0.2111***	–0.2109***	–0.0668	–0.0436	–0.0727	–0.0502	–0.0626	–0.0413	–0.0737*	–0.0489
	[0.0381]	[0.0374]	[0.0491]	[0.0475]	[0.0468]	[0.0452]	[0.0447]	[0.0429]	[0.0429]	[0.0411]
Shock, nonusers | user Xs							–0.1230**	–0.1219**	–0.1024**	–0.1010**
							[0.0549]	[0.0548]	[0.0502]	[0.0500]
Mean of user	0.5656	0.5663	0.5656	0.5663	0.5656	0.5663	0.5661	0.5668	0.5661	0.5668

Notes: Dependent variable is log total household consumption per capita. Heteroscedasticity-robust standard errors are in brackets. Controls: household demographics; household head education and occupation; use of bank accounts, SACCOs and ROSCAs; cell phone ownership. Interactions refer to interactions of the controls with the shock. When interactions are included, the overall effect of shock is evaluated at the mean of the covariates. ***, **, * Significant at 1, 5 and 10 per cent levels.

3.2 Propensity score matching

The main challenge of a credible impact evaluation is to construct counterfactual outcomes; that is, identifying what would have happened to M-PESA users in the absence of M-PESA. Propensity score matching was developed to help design and analyse non-randomized observational studies so that they mimic some of the characteristics of a randomised controlled trial. It allowed us to account for the possible differences in wealth, education and other socioeconomic characteristics between M-PESA users and nonusers.

As Table 6 shows, M-PESA users are wealthier, more educated and more likely to own a cell phone, and have better access to formal financial services, such as bank account.

Table 6: Summary statistics by user status

	M-PESA users		M-PESA nonusers	
	Mean	SD	Mean	SD
Own cell phone	0.917	0.276	0.473	0.499
Per capita consumption	83,754	127,791	48,532	81,079
Per capita food consumption	34,659	29,192	26,117	26,877
Total wealth	183,094	729,928	67,367	259,785
HH size	4.338	2.229	4.349	2.336
Education of head (years)	8.313	5.321	5.872	5.079
Positive shock	0.092	0.289	0.082	0.274
Negative shock	0.533	0.499	0.539	0.499
Weather/Agricultural shock	0.086	0.280	0.086	0.281
Illness shock	0.344	0.475	0.297	0.457
Send remittances	0.614	0.487	0.266	0.442
Receive remittances	0.538	0.499	0.229	0.420
Financial access dummies				
Bank account	0.669	0.471	0.299	0.458
Mattress	0.710	0.454	0.813	0.390
Savings and credit cooperative	0.222	0.416	0.130	0.337
Merry go round/ROSCA	0.471	0.499	0.381	0.486
Household head occupation dummies				
Farmer	0.197	0.398	0.390	0.488
Public service	0.049	0.216	0.016	0.126
Professional occupation	0.249	0.432	0.169	0.374
Househelp	0.112	0.316	0.079	0.270
Run a business	0.160	0.367	0.145	0.353
Sales	0.087	0.282	0.048	0.214
In industry	0.023	0.151	0.028	0.165
Other occupation	0.049	0.216	0.054	0.226
Unemployed	0.070	0.256	0.069	0.253
Number of observations	2,773		1,791	

Matching involves constructing a new control group using background characteristics, so that for every treated observation (M-PESA user), there is an untreated one (nonuser) that is as similar as possible in observable characteristics. Propensity score matching constructs a probability that a household adopts M-PESA conditional on its characteristics. This is done by running a binary probit regression of M-PESA users and nonusers on the set of observable baseline characteristics (in particular, household demographics, years of education of the household head, occupations for the household head, the use of bank accounts, the use of savings and credit cooperatives, the use of rotating savings and credit associations, and cell phone ownership). It can be written as:

$$P(x) = Pro\,b[T = 1|X = x_1, x_2, \dots, x_k] = E(T|X), \tag{5}$$

where $T = 1$ for households using M-PESA and 0 otherwise, and X is a vector of characteristics.

To investigate the validity of the propensity score matching estimation, I need to verify the common support or overlap condition. This condition requires the existence of a sufficient overlap between the propensity scores of treated and control groups. Through the following histogram (Figure 2) and density distribution plots (Figure 3), I am able to visually inspect whether the matching is able to make the distributions more similar. For the matching, I use the nearest-neighbour matching method, in which each M-PESA user is matched to its nearest neighbour nonuser based on the propensity score. Performing 1:1 nearest-neighbour matching results in keeping, on average, 70 per cent of the sample. Increasing the matching ratio could increase precision but also bias.

Figure 2: Histograms of matched sub-samples along common support

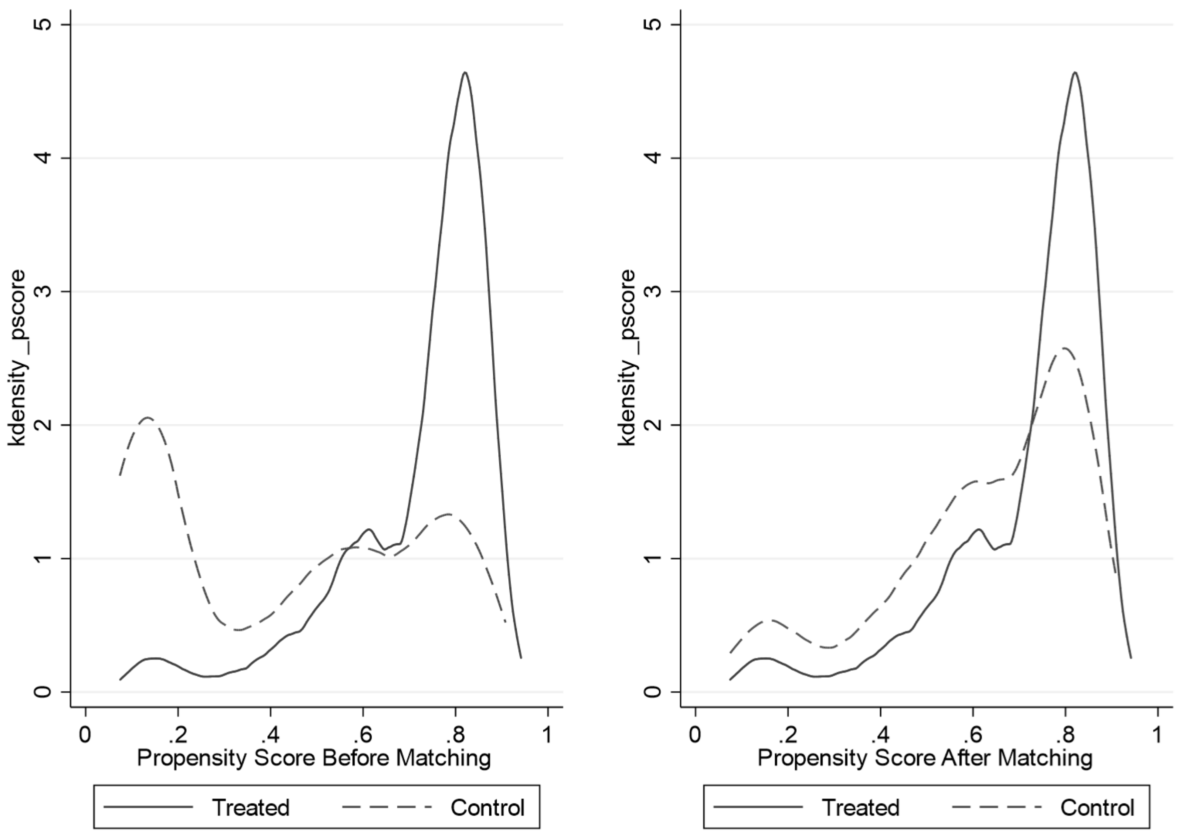
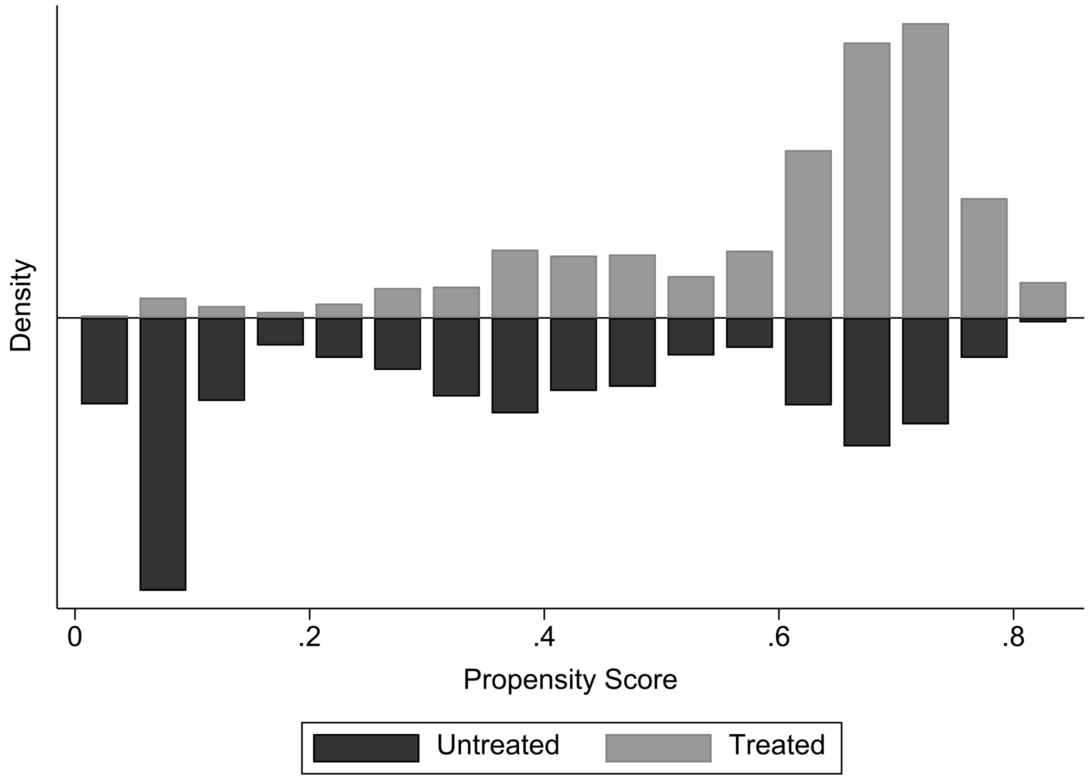


Figure 3: Estimated probability of being an M-PESA user (propensity score)



First, the probability of being an M-PESA user, given the values of all potential confounders (the propensity score, PS, of being a user), was estimated for each household in the sample. This was done by estimating a probit regression model with all known confounders as predictors. Then, each user was 'matched' with one nonuser, based on similar values of the propensity scores. This allowed construction of a control group as similar as possible to the treatment group of M-PESA users. In this way, propensity score matching tried to mimic random assignment of households across treatment and control groups.

Analytic weights were employed to account for the fact that some nonuser households were matched to more than one user household. I then estimated the specification in equation (3).

Table 7 reports results of the basic specifications. From column (1), I see that a negative shock reduces per capita consumption of nonuser households by 21 per cent. This is consistent with the results reported in Table 4A. In contrast, households with M-PESA are better able to protect themselves against negative shocks. M-PESA users cut back consumption in the face of a negative shock by only 6 per cent (–0.2127 + 0.1505). This figure is approximately half the size of the original paper but still significant. According to the last column, in which a full set of controls is included, nonusers suffer from approximately 15 per cent reduction in consumption (the corresponding figure is 7 per cent in Table 4A), whereas users are able to smooth their consumption perfectly and experience no significant reduction in consumption. Across Table 7, the coefficient on the interaction of interest is strongly significant. The results are robust to the new weighting system driven from propensity score matching.

Table 7: Basic difference-in-differences results, propensity score matching (nearest neighbour matching)

	Total consumption				
	OLS	Panel	Panel	Panel	Panel
	(1)	(2)	(3)	(4)	(5)
M-PESA user	0.1368***	–0.0311	–0.0210	–0.0251	–0.0329
	[0.0465]	[0.0469]	[0.0461]	[0.0521]	[0.0458]
Negative shock	–0.2127***	–0.1344**	–0.1244**	0.2603	0.2895
	[0.0507]	[0.0600]	[0.0587]	[0.1989]	[0.1926]
User × negative shock	0.1505**	0.1689**	0.1673**	0.2199***	0.1908***
	[0.0624]	[0.0699]	[0.0696]	[0.0816]	[0.0703]
Demographic controls			Yes	Yes	Yes
Controls + interactions				Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Time × location FE		Yes	Yes		Yes
Observations	3,169	3,169	3,169	3,159	3,159
Negative shock	–0.1132***	–0.0228	–0.0139	–0.0171	–0.0201
	[0.0296]	[0.0298]	[0.0286]	[0.0290]	[0.0284]
Shock, users	–0.0621*	0.0345	0.0429	0.0573	0.0443
	[0.0364]	[0.0339]	[0.0334]	[0.0351]	[0.0329]
Shock, nonusers	–0.2127***	–0.1344**	–0.1244**	–0.1624**	–0.1457**
	[0.0507]	[0.0600]	[0.0587]	[0.0643]	[0.0578]
Shock, nonusers | user Xs				–0.1626**	–0.1465**
				[0.0672]	[0.0597]
Mean of user	0.6605	0.6605	0.6605	0.6613	0.6613

Notes: Dependent variable is log total household consumption per capita. Heteroscedasticity-robust standard errors are in brackets. Controls: household demographics; household education and occupation; use of bank accounts, SACCOs and ROSCAs; cell phone ownership. Interactions refer to interactions of the controls with the shock. When interactions are included, the overall effect of shock is evaluated at the mean of the covariates. ***, **, * Significant at 1, 5 and 10 per cent levels.

3.3 Tobit model

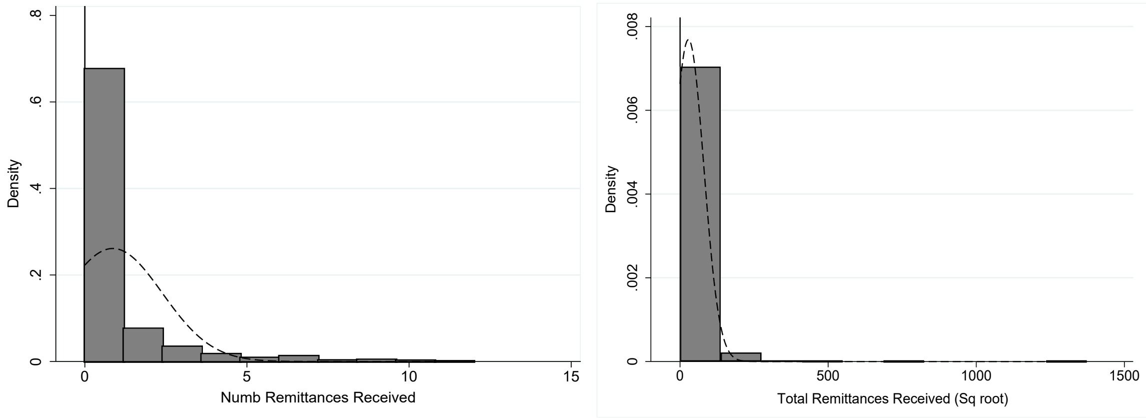
According to Jack and Suri (2014), remittances are the main channel through which households can share risks in the event of any negative shocks. The role of remittances is examined by estimating the following specification, reported as equation (8) in the original paper:

$$r_{ijt} = \alpha_i + \gamma Shock_{ijt} + \mu User_{ijt} + \beta User_{ijt} \times Shock_{ijt} + \theta X_{ijt} + \eta_{jt} + \pi_{rt} + \varepsilon_{ijt}, \quad (6)$$

where the dependent variable, r_{ijt}, is a measure of remittances over the previous six months. This variable can be measured as (i) the probability of receiving remittances, (ii) the number of remittances received or (iii) the total value of remittances received. It is worth mentioning that the last two outcome variables are not available (or, in other words, take the value of zero) for those households who never received any remittances.

Due to the censored nature of these two measures of remittances, common linear regressions may produce biased estimates. The following graphs clearly demonstrate the frequency of zero observations (Figure 4, sections A and B).[14] Given the zeros for nonusers, these variables have a skewed distribution.

Figure 4: Number of remittances received (A) and total value received (B)



(A) (B)

To avoid bias stemming from left-censoring in the dependent variable, I employ a Tobit model, also called a censored regression model. Table 8 and Table 9 report the Tobit estimates of the number and total value of remittances received. Across the table, the interaction of interest is positive and significant, meaning that in the event of a negative shock, users receive more remittances in terms of the number received and total value. This conclusion is consistent with the corresponding fixed effect model estimates from Table 5A.

In fact, controlling for the censored nature of remittance variables increases the size of coefficients on the interaction term.[15] In Table 8 and Table 9, I report the original findings on the grey-highlighted columns. The next two columns report multilevel mixed effects Tobit regressions, with and without left-censoring limit.

[14] Due to the large frequency of zeros compared to the other values, the bar charts in the right-hand side of Figure 4 (B) are not obvious here.

[15] The reason why I also report the results corresponded to the Tobit without controlling for left-censoring is to show that the estimated effects are not far from the original results, once the left-censoring issue is taken into account. It is worth mentioning that these estimates are not particularly large if they are compared to the average size: the mean values for 'Number received' and 'Total received (square root)' are 2.067 and 68.50, respectively.

Table 8: Mechanisms (panel), Tobit model

	Overall shock: sample without Nairobi					
	Number received			Total received (square root)		
	Original (1)	Tobit without ll (0) (2)	Tobit with ll (0) (3)	Original (4)	Tobit without ll (0) (5)	Tobit with ll (0) (6)
M-PESA user	0.2574**	0.5181***	1.675**	10.6757***	19.8879**	52.1128**
	[0.1305]	[0.0938]	[0.2418]	[3.7863]	[2.5962]	[6.5803]
Negative shock	–0.1306	0.0419	–0.7118	1.8775	4.3335	–27.7752
	[0.4193]	[0.3323]	[0.7321]	[12.0864]	[8.1873]	[18.5291]
User × shock	0.3286*	0.3648***	0.8502***	8.3428*	5.4418*	14.2875*
	[0.1789]	[0.1242]	[0.3156]	[4.6884]	[3.2583]	[8.3201]
Controls + interactions	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Time × location FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,911	3,911	3,911	3,873	3,873	3,873
Shock effect	–0.0369	0.0582	0.0412	1.6647	1.3617	1.1966
	[0.0871]	[0.0670]	[0.1456]	[2.2697]	[1.6822]	[3.8094]
Shock, users	0.0470	0.1680*	0.3172*	4.3755	3.4506	6.5569
	[0.1157]	[0.0923]	[0.1672]	[3.4195]	[2.5122]	[4.4885]
Shock, nonusers	–0.1400	–0.0767	–0.2978	–1.6403	–1.1850	–5.3388
	[0.1221]	[0.0825]	[0.2325]	[2.6656]	[1.8193]	[5.9558]
Mean of user	0.5512	0.5512	0.5512	0.5494	0.5494	0.5494

Notes: Replication results applying Tobit models with and without left-censoring limit are in column 2, 3, 5, 6, 8, 9, 11 and 12. Original results are in columns 1, 4, 7 and 10. ***, **, * Significant at 1, 5 and 10 per cent levels.

Table 9: Mechanisms (panel), Tobit model (continued)

	Overall shock: without Mombasa			Illness shock		
	Total received (square root)			Total received (square root)		
	Original (7)	Tobit without ll (0) (8)	Tobit with ll (0) (9)	Original (10)	Tobit without ll (0) (11)	Tobit with ll (0) (12)
M-PESA user	9.0579**	20.1995***	50.4186***	12.5548***	21.0130***	55.3657***
	[4.0683]	[2.7787]	[6.8415]	[3.1596]	[2.2776]	[5.7112]
Negative shock	–1.8885	3.6285	–31.7009*	–9.3597	–4.5909	–41.0790**
	[12.4371]	[8.2745]	[18.4973]	[10.9683]	[8.1705]	[19.6354]
User × shock	10.0472**	5.2337	14.8716*	8.6003	5.8520	14.0697
	[4.9200]	[3.4071]	[8.5149]	[5.2788]	[3.6269]	[8.8802]
Controls + interactions	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Time × location FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,665	3,665	3,665	3,873	3,873	3,873
Shock effect	1.5026	1.4530	1.3329	2.7412	1.4099	0.6455
	[2.3569]	[1.7380]	[3.8885]	[2.5233]	[1.8090]	[4.0562]
Shock, users	4.6901	3.4386	6.9130	6.5410*	4.3611	7.3552
	[3.5678]	[2.6197]	[4.6184]	[3.5215]	[2.6556]	[4.7119]
Shock, nonusers	–2.3154	–0.9255	–5.3512	–1.8914	–2.1881	–7.5350
	[2.7528]	[1.8810]	[6.0308]	[3.0544]	[2.0793]	[6.5027]
Mean of user	0.5450	0.5450	0.5450	0.5494	0.5494	0.5494

Notes: Replication results applying Tobit models with and without left-censoring limit are in column 2, 3, 5, 6, 8, 9, 11 and 12. Original results are in columns 1, 4, 7 and 10. ***, **, * Significant at 1, 5 and 10 per cent levels.

3.4 Endogeneity and IV regressions

Jack and Suri provide a discussion of potential sources of endogeneity in the basic specification (equation 7 in the paper):

$$C_{ijt} = \alpha_i + \gamma Shock_{ijt} + \mu User_{ijt} + \beta User_{ijt} \times Shock_{ijt} + \theta^S X_{ijt} \times Shock_{ijt} + \theta X_{ijt} + \eta_{jt} + \pi_{rt} + \varepsilon_{ijt} \quad (7)$$

They then explain that, to identify the causal effect of M-PESA on risk sharing, I must assume the interaction term $User_{ijt} \times Shock_{ijt}$ is exogenous or uncorrelated with error term.

3.4.1 Endogeneity of shocks

The authors emphasize that their identification assumption is satisfied if shocks are truly exogenous. They argue that this may be reasonable, for two reasons: (i) households were asked in the survey to report only unexpected events that affected them; and (ii) reported shocks are not systematically correlated with a number of household-level variables.

Income shocks are correlated with consumption changes and remittances, as would be expected, but they are not correlated with other household characteristics, nor with access to agents or M-PESA use. They report these correlations in Table 3, page 202.

3.4.2 Endogeneity of M-PESA use

The authors also argue that any endogeneity of M-PESA use due to selective adoption associated with wealth or other unobservables should be absorbed in the main effect of being a user (self-selection effects into using M-PESA are absorbed into coefficient μ on $User_{ijt}$). Although they do not report results relating M-PESA usage to household characteristics, they control for these by adding X_{ijt} (a vector of controls – in particular, household demographics, household head years of education and occupation dummies, the use of financial instruments and a dummy for cell phone ownership) in all the regressions. They also control for unobservable characteristics of households by using fixed effects. They argue that this approach allows for unobservables to be correlated with M-PESA use, as long as those unobservables do not interact with the response to shock.

However, fixed effects do not correct all sources of endogeneity. For instance, people who are especially innovative might be more willing to adopt M-PESA than their counterparts, while choosing strategies that minimize their exposure to shocks. Alternatively, people who live in places that experience more shocks might be more willing to adopt M-PESA. In such cases, estimates of the interaction coefficient for shocks and M-PESA use would be biased.

To deal with endogeneity concerns, Jack and Suri propose two strategies: (i) they extend equation (6) to include interactions of the shock variable with all observable covariates; and (ii) they use agent rollout data and incorporate that in a standard IV procedure (equation (4)). They argue that the agent variable (geographic proximity to the agents as indicators of access) is exogenous.

This latter statement is debatable. There are reasons to support that M-PESA agents are not exogenously distributed across geographical areas. M-PESA agents receive a commission on a sliding scale for deposits and withdrawals but not for transactions. Thus, there is an incentive for M-PESA agents to locate where they can gain the greatest profit. It seems likely that wealthier areas would provide greater profit opportunities, because there is the potential for larger sums of money to be handled.

I first address this concern by controlling for the interactions between observable individual characteristics and shocks. The specification would then be as follows:

$$C_{ijt} = \alpha_i + \gamma Shock_{ijt} + \vartheta Agent_{ijt} + \beta Agent_{ijt} \times Shock_{ijt} + \theta^S X_{ijt} \times Shock_{ijt} + \theta^M X_{ijt} + \eta_{jt} + \pi_{rt} + \varepsilon_{ijt} \quad (8)$$

The results are presented in Table 10. As can be seen, whenever demographic controls are available, I also control for the interaction of the shock with all observable covariates. This means I am not adding this control factor for column 2 (the grey-highlighted column). Of the eight coefficients on the interaction of interest, the first six remain statistically significant.

However, when I test for weak instruments, I find evidence of weak instruments in every specification other than the original paper's specification (column 2). For the Wald test, when there are two endogenous variables and four instruments, Stock and Yogo (2005) suggest a critical value of 11.[16] In every specification other than column 2, the Kleibergen–Paap test statistic is less than the critical value, indicating the instruments inadequately address bias and size issues.[17]

Table 10: IV results (cross-section and panel), adding demographic controls

	Total consumption					Pr [Receive]		
	Cross-section	Panel	Panel	Panel without Mombasa	Panel	Panel without Mombasa	Panel without Mombasa	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
M-PESA user	–0.9786**	–0.5128***	–0.8740***	–1.0184***	–0.4303	–0.5649	–0.2991	–0.1070
	[0.4582]	[0.1885]	[0.2954]	[0.3481]	[0.8047]	[1.3868]	[0.2824]	[1.0379]
Negative shock	–0.3244*	–0.3601**	0.0659	–0.0094	0.0720	0.0002	–0.0925	–0.1076
	[0.1689]	[0.1668]	[0.1907]	[0.2210]	[0.1804]	[0.2090]	[0.1654]	[0.1578]
User × shock	1.5627**	0.6146**	1.1620**	1.4717**	1.3529**	1.5361**	0.6694	0.5832
	[0.6887]	[0.2908]	[0.5402]	[0.6081]	[0.5463]	[0.6369]	[0.5048]	[0.5000]
Controls interactions	**Yes**		**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
Controls	Yes		Yes	Yes	Yes	Yes	Yes	Yes
Rural × time FE					Yes	Yes		Yes
Location + rural FE	Yes							
Observations	3,911	3,926	3,894	3,688	3,894	3,688	3,688	3,688
Shock effect	0.5371	–0.0217	0.7064**	0.7956***	0.8178***	0.8404***	0.2736	0.2114
	[0.3422]	[0.0365]	[0.2831]	[0.2947]	[0.2899]	[0.3017]	[0.2488]	[0.2486]
Shock, users	1.2384	0.2545*	1.2279**	1.4622***	1.4249***	1.5363***	0.5768	0.4756
	[0.6340]	[0.1331]	[0.4968]	[0.5346]	[0.5098]	[0.5603]	[0.4537]	[0.4534]
Shock, nonusers	–0.3244**	–0.3601**	0.0659	–0.0094	0.0720	0.0002	–0.0925	–0.1076
	[0.1689]	[0.1668]	[0.1907]	[0.2210]	[0.1804]	[0.2090]	[0.1654]	[0.1578]
Kleibergen–Paap LM test	23.000	46.681	22.698	24.163	4.015	2.192	24.163	2.192
LM test *p*-value	0.0000	0.0000	0.0000	0.0000	0.2598	0.5335	0.0000	0.5335
Kleibergen–Paap *F* statistic (critical value = 11)	5.670	12.346	5.226	6.319	0.986	0.533	6.319	0.533

Notes: Heteroscedasticity-robust standard errors in brackets. ***, **, * Significant at 1, 5 and 10 per cent levels.

[16] The Cragg and Donald (1993) statistics can be used to assess the overall strength of the instruments, and Stock and colleagues (2005) have tabulated critical values of the minimum eigenvalue of the Cragg–Donald statistics for testing whether the instrument is weak. This critical value (which is equal to 11) is reported in brackets for the comparison.

[17] Like Jack and Suri (2014), I also report the F-stat form of the Kleibergen–Paap statistic, the heteroskedastic and clustering robust version of the Cragg–Donald statistic suggested by Stock and colleagues (2005) as a test for weak instruments. All statistics are well below the critical value (except the second column), confirming the possibility of weak instrument.

In the next step, I report results for some of the first-stage regressions. Although Jack and Suri do not provide the associated first-stage results in their paper, they do indicate that predicting M-PESA use is not very precise in some cases. Table 11 reports my results. For space reasons, I show the first-stage results for cross-section and the most comprehensive panel (column 5, above). I also report the first-stage regressions when the outcome variable is a measure of remittances.

I then evaluate whether the instrument applied in the original paper is weak. The proposed instruments for the use of M-PESA and their interactions with shock according to paper are as follows: (i) distance to the closest agent, (ii) the number of agents within 5 kilometres of the household and (iii) the interaction of each with the shock.

As observed from the first-stage results, distance to the closest agent and the number of agents within 5 kilometres of the household are not very strong determinants of M-PESA adoption. This is confirmed by small first-stage F-statistic. The Kleibergen–Paap F-statistics for the weak instrument test are not always larger than critical value for the Stock–Yogo, which creates the possibility of weak instrument problem.

As noted in my earlier discussion about legitimacy of the chosen instruments, I planned to replace the instrument with potential alternatives. My planned approach was to find an instrument (or a set of instruments) correlated with M-PESA use (relevant) but not affecting consumption smoothing (exogenous). In the literature, Mbiti and Weil (2011) apply an alternative instrument for M-PESA use in Kenya. They focus on the 2006 perception data (before the advent of M-PESA in 2007) and explain that the survey included a question referring to the perceptions of the most common money transfer methods. The households were then asked to identify the slowest, riskiest and most costly transfer methods. They conclude that those households who felt their means of transfer was not efficient were more likely to adopt M-PESA after its introduction. I checked the survey instrument (questionnaire) and codebook to see whether such a perception question was available. It turns out that there is a question asking about the method used to send money, followed by a question on the reason why the method was used (easy, cheap, safe, fast and other). Unfortunately, I was unable to construct this instrument, due to the unavailability of data.

The combination of (i) concern over weak instruments and (ii) characteristics differences between M-PESA users and nonusers serves to weaken the strength of Jack and Suri's arguments.

Table 11: IV results (cross-section and panel), first-stage estimations

	Total consumption					
	Cross section IV–2nd	Cross section IV–1st	Cross section IV–1st	Panel IV–2nd	Panel IV–1st	Panel IV–1st
	(1)	(2)	(3)	(4)	(5)	(6)
		M-PESA user	User × shock		M-PESA user	User × shock
M-PESA user	–0.4705*			–0.3155		
	[0.2685]			[0.8549]		
Negative shock	–0.3344**			–0.3762**		
	[0.1469]			[0.1547]		
User × shock	0.5124*			0.6782**		
	[0.2653]			[0.2679]		
Dis to closest agent		0.0114 [0.0122]	0.0063 [0.0070]		0.0038 [0.0280]	–0.0270 [0.0239]
Number of agents in 5 km		0.0376*** [0.0063]	0.0103*** [0.0038]		0.0229* [0.0129]	–0.0066 [0.0094]
Dis × shock		–0.0352***	–0.0551***		–0.0156	–0.0391***
		[0.0130]	[0.0095]		[0.0162]	[0.0132]
Agent 5 × shock		–0.0110*	0.0112**		–0.0059	0.0195***
		[0.0064]	[0.0049]		[0.0082]	[0.0072]
Negative shock		0.3156***	0.9360***		0.1185	0.7920***
		[0.1085]	[0.0790]		[0.1385]	[0.1110]
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Rural × time FE				Yes	Yes	Yes
Location + rural FE	Yes	Yes	Yes			
Observations	3,911	3,911	3,911	3,894	3,894	3,894
Shock effect	–0.0519**			–0.0024		
	[0.0264]			[0.0314]		
Shock, users	0.1781			0.3020**		
	[0.1237]			[0.1205]		
Shock, nonusers	–0.3344** [0.1469]			–0.3762** [0.1547]		
Kleibergen–Paap LM test	39.796			3.7324		
LM test *p*-value	0.0000			0.2919		
Kleibergen–Paap *F* statistic (critical value =11)	9.390			0.9262		

Notes: Heteroscedasticity-robust standard errors in brackets. ***, **, * Significant at 1, 5 and 10 per cent levels.

Table 12: IV results (cross-section and panel), first-stage estimations (continued)

	Pr{Receive}					
	Panel without Msa IV–2nd (7)	Panel without Msa IV–1st (8)	Panel without Msa IV–1st (9)	Panel without Msa IV–2nd (10)	Panel without Msa IV–1st (11)	Panel without Msa IV–1st (12)
		M-PESA user	User × shock		M-PESA user	User × shock
M-PESA user	–0.1455			0.0849		
	[0.1741]			[1.1219]		
Negative shock	–0.1739			–0.1395		
	[0.1584]			[0.1718]		
User × shock	0.3862			0.3080		
	[0.2732]			[0.2903]		
Dis to closest agent		–0.0114	–0.0328		–0.0027	–0.0283
		[0.0299]	[0.0251]		[0.0288]	[0.0248]
Number of agents in 5 km		0.1526***	0.0568***		0.0214	–0.0046
		[0.0137]	[0.0101]		[0.0199]	[0.0162]
Dis × shock		–0.0273	–0.0430***		–0.0132	–0.0350**
		[0.0179]	[0.0143]		[0.0169]	[0.0140]
Agent 5 × shock		–0.0021	0.0299***		–0.0040	0.0287***
		[0.0122]	[0.0104]		[0.0114]	[0.0099]
Negative shock		0.2017	0.8096		0.0977	0.7487***
		[0.1536]	[0.1231]		[0.1476]	[0.1216]
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Rural × time FE				Yes	Yes	Yes
Location + rural FE						
Observations	3,688	3,688	3,688	3,688	3,688	3,688
Shock effect	0.0373			0.0290		
	[0.0302]			[0.0313]		
Shock, users	0.2123*			0.1685		
	[0.1212]			[0.1246]		
Shock, nonusers	–0.1739			–0.1395		
	[0.1584]			[0.1718]		
Kleibergen–Paap LM test	45.2436			1.7344		
LM test *p*-value	0.0000			0.6293		
Kleibergen–Paap *F* statistic (critical value =11)	12.5678			0.4258		

Notes: Msa refers to Mombasa. Heteroscedasticity-robust standard errors in brackets. ***, **, * Significant at 1, 5 and 10 per cent levels.

3.5 Longer period balanced and unbalanced panel

Jack and Suri conducted five surveys across Kenya between 2008 and 2014. These data sets are publicly available through the FSD Kenya website.[18] Since the original study made use of only the first three rounds (or two periods), my provisional plan was to extend the years of analysis to the most recent available (see Table 13). Unfortunately, data from round 4 is not publicly available. Despite data availability for round 5, I was unable to match households over time.[19] In addition, as the authors mention in the 'Data Read Me' file, some variables that contain personally identifiable information are dropped from the data sets. As a result, this plan could not be completed. Instead, I dropped 265 households corresponding to the third round. Therefore, no further control for the difference in the timing of the survey between rounds 2 and 3 is required.

Table 13: Survey time frame

	Time frame	Sample households	Number of data sets
Round 1	August 2008–October 2008	3,000	23
Round 2	October 2009–January 2010	2,105	21
Round 3	May 2010–August 2010	1,531	24
Round 4	March 2011–June 2011	1,649	Not available
Round 5	June 2014–September 2014	1,688	19

Jack and Suri provide a description of their survey data in Section III of their paper. They explain that the number of households surveyed in the first round (September 2008) was 3,000. However, they were not able to find the original (full) sample in the follow-up surveys conducted in December 2009 and June 2010. As a result, the total number of households interviewed in 2009 was 2,017. This number dropped to 1,595 in the following year (2010), 265 of whom were not interviewed in 2009. In other words, there were only 1,330 (1,595 – 265 = 1,330) common among the three rounds.[20] This means an attrition rate of about 56 per cent. In order to reduce the high attrition rate, the authors add two panels of 2,017 (from round 2) and 265 (from round 3) together, which are lumped together as a 'second period'. This strategy allows them to construct a balanced, two-period panel of 2,282 households. They distinguish the time differences in period 2 by including a time fixed effect.[21]

To check for robustness, I restricted my analysis to a sample comprising those interviewed in the first two rounds. This allowed me to construct a balanced, two-period panel of 2,017 households in lieu of 2,282. Table 14 and Table 15 report the results of basic specifications presented in Table 4A of Jack and Suri's paper, while using a restricted sample. Since the second period refers only to the second round (it previously referred to the second and third rounds), I do not control for round (time) dummies

[18] The data is publicly available at: http://fsdkenya.org/dataset/m-pesa-panel-survey-kenya-2014/
[19] I checked some demographics – such as gender, which is fixed over time – to see whether households could be matched through the unique household identifier (hhid) provided for all the subsets. As it turns out, it could not be matched. This was also confirmed through private correspondence with Professor Suri.
[20] According to footnote 20, the three-period balanced panel includes 1,311 households.
[21] Period 2 involves rounds 2 and 3.

throughout the regression analysis presented in Table 15. On the other hand, Table 14 reports the results for this restricted sample, while still including time (round) dummies.

Table 14: Basic difference-in-differences results, restricted sample

	Total consumption Full sample				
	OLS	Panel	Panel	Panel	Panel
	(1)	(2)	(3)	(4)	(5)
M-PESA user	0.5556***	0.0506	0.0424	–0.0260	–0.0077
	[0.0400]	[0.0501]	[0.0488]	[0.0502]	[0.0468]
Negative shock	–0.2112***	–0.0683	–0.0775	0.2293	0.2636
	[0.0402]	[0.0514]	[0.0481]	[0.2043]	[0.1909]
User × negative shock	0.0915*	0.1162*	0.1465**	0.1782***	0.1541**
	[0.0537]	[0.0639]	[0.0594]	[0.0672]	[0.0617]
Demographic controls			Yes	Yes	Yes
Controls + interactions				Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Time × location FE		Yes	Yes		Yes
Observations	4,034	4,034	4,034	4,019	4,019
Negative shock	–0.1607***	–0.0033	0.0045	0.0076	–0.0022
	[0.0268]	[0.0318]	[0.0303]	[0.0311]	[0.0293]
Shock, users	–0.1204***	0.0479	0.0690*	0.0630	0.0531
	[0.0358]	[0.0391]	[0.0378]	[0.0411]	[0.0363]
Shock, nonusers	–0.2119***	–0.0682	–0.0775	–0.0629	–0.0725*
	[0.0402]	[0.0514]	[0.0481]	[0.0469]	[0.0452]
Shock, nonusers | user Xs				–0.1152**	–0.1010*
				[0.0567]	[0.0522]
Mean of user	0.5595	0.5595	0.5595	0.5600	0.5600

Notes: Dependent variable is log total household consumption per capita. Heteroscedasticity-robust standard errors are reported in brackets. Controls: household demographics; household education and occupation; use of bank accounts, SACCOs and ROSCAs; cell phone ownership. Interactions refer to interactions of the controls with the shock. When interactions are included, the overall effect of shock is evaluated at the mean of the covariates. ***, **, * Significant at 1, 5 and 10 per cent levels.

According to the baseline results in Table 15, the last two columns are identical. This means that once I am not controlling for time – and therefore its interaction with location and rural areas – these additional dummies have no impact on my results. Although the general conclusions remain the same, the first column shows that in the event of negative shocks, per capita consumption for users falls by 15 per cent. The corresponding estimate from the full sample in Table 4A is 12 per cent.[22]

[22] The OLS results are reported as a baseline for the sake of comparison. It does not include any controls except time fixed effects.

Table 15: Basic difference-in-differences results, restricted sample

	Total consumption				
	Full sample				
	OLS	Panel	Panel	Panel	Panel
	(1)	(2)	(3)	(4)	(5)
M-PESA user	0.5108***	–0.0149	–0.0243	–0.0965**	–0.0965**
	[0.0398]	[0.0521]	[0.0512]	[0.0491]	[0.0491]
Negative shock	–0.2240***	–0.0662	–0.0706	0.2349	0.2349
	[0.0404]	[0.0529]	[0.0502]	[0.2128]	[0.2128]
User × negative shock	0.0789*	0.0942	0.1111	0.1654**	0.1654**
	[0.0540]	[0.0719]	[0.0698]	[0.0685]	[0.0685]
Demographic controls			Yes	Yes	Yes
Controls + interactions				Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Location FE		Yes	Yes		Yes
Observations	4,034	4,034	4,034	4,019	4,019
Negative shock	–0.1798***	–0.0134	–0.0084	–0.0099	–0.0099
	[0.0268]	[0.0349]	[0.0340]	[0.0317]	[0.0317]
Shock, users	–0.1450***	0.0280	0.0406	0.0370	0.0370
	[0.0358]	[0.0474]	[0.0472]	[0.0422]	[0.0422]
Shock, nonusers	–0.2240***	–0.0661	–0.0706	–0.0695	–0.0695
	[0.0404]	[0.0529]	[0.0502]	[0.0471]	[0.0471]
Shock, nonusers | user Xs				–0.1284**	–0.1284**
				[0.0580]	[0.0580]
Mean of user	0.5595	0.5595	0.5595	0.5600	0.5600

Notes: Dependent variable is log total household consumption per capita. Heteroscedasticity-robust standard errors are reported in brackets. Controls: household demographics; household education and occupation; use of bank accounts, SACCOs and ROSCAs; cell phone ownership. Interactions refer to interactions of the controls with the shock. When interactions are included, the overall effect of shock is evaluated at the mean of the covariates. ***, **, * Significant at 1, 5 and 10 per cent levels.

3.6 Multilevel fixed effects with multiway clustering

Finally, I conducted an additional robustness check to examine whether the main results are robust to an alternative estimation that considers multilevel fixed effects – namely households, locations and time fixed effects. I further included two-way clustering (households and locations). The corresponding results are reported in Table 16. As can be seen, these estimates are consistent with the original findings.

Table 16: Basic difference-in-differences results, multilevel fixed effects with multiway clustering

	Total consumption Full sample							
	Panel (2)		Panel (3)		Panel (4)		Panel (5)	
	Original panel (1)	MFE-MC panel (2)	Original panel (3)	MFE-MC panel (4)	Original panel (5)	MFE-MC panel (6)	Original panel (7)	MFE-MC panel (6)
M-PESA user	0.0520	0.0520	0.0456	0.0456	–0.0223	–0.0223	–0.0088	–0.0088
	[0.0481]	[0.0487]	[0.0469]	[0.0556]	[0.0484]	[0.0558]	[0.0449]	[0.0498]
Negative shock	–0.0688	–0.0668	–0.0727	–0.0727	0.2872	0.2872	0.2673	0.2673
	[0.0491]	[0.0497]	[0.0468]	[0.0596]	[0.1762]	[0.1673]	[0.1799]	[0.1846]
User × shock	0.1093*	0.1093*	0.1320**	0.1320*	0.1749***	0.1749***	0.1483**	0.1483**
	[0.0616]	[0.0623]	[0.0594]	[0.0728]	[0.0663]	[0.0699]	[0.0599]	[0.0661]
Demographic controls			Yes	Yes	Yes	Yes	Yes	Yes
Controls + Interactions					Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time × location FE	Yes	Yes	Yes	Yes			Yes	Yes
Observations	4,562	4,500	4,562	4,500	4,545	4,526	4,545	4,466
Negative Shock	–0.0050	–0.0052	0.0019	0.0016	0.0022	0.0022	–0.0059	–0.0065
	[0.0305]	[0.0308]	[0.0292]	[0.0330]	[0.0286]	[0.0337]	[0.0280]	[0.0323]
Shock, users	0.0425	0.0425	0.0592	0.0592	0.0518	0.0515	0.0460	0.0452
	[0.0379]	[0.0383]	[0.0370]	[0.0386]	[0.0383]	[0.0429]	[0.0355]	[0.0374]
Shock, nonusers	–0.0668	–0.0668	–0.0727	–0.0727	–0.0626	–0.0623	–0.0737*	–0.0734
	[0.0491]	[0.0497]	[0.0468]	[0.0596]	[0.0447]	[0.0550]	[0.0429]	[0.0556]
Shock, nonusers | user Xs					–0.1230**	–0.1233**	–0.1024**	–0.1031*
					[0.0549]	[0.0614]	[0.0502]	[0.0592]
Mean of user	0.5656	0.5634	0.5656	0.5634	0.5661	0.5666	0.5661	0.5643

Notes: MFE-MC refers to multilevel fixed effects, multiway clustering. Dependent variable is log total household consumption per capita. Heteroscedasticity-robust standard errors are reported in brackets. Controls: household demographics; household education and occupation; use of bank accounts, SACCOs and ROSCAs; cell phone ownership. Interactions refer to interactions of the controls with the shock. When interactions are included, the overall effect of shock is evaluated at the mean of the covariates. ***, **, * significant at 1, 5 and 10 per cent levels.

4. Theory of change analysis

Thus far, it is clear that the results in Jack and Suri (2014) are robust to alternative measurement and estimation analysis, confirming that M-PESA users, compared to nonusers, are better able to protect their consumption in the event of a negative shock. This is mainly attributed to associated reduction in transaction costs. The remaining question is whether the effect is heterogeneous, depending on the rural–urban status. In this section, I shift the focus of analysis to answer this question.

The main focus of the Jack and Suri (2014) study is Kenyan households across a large part of the country, regardless of their residence locations (rural or urban area).[23] According to the discussion in the original study, urban households have higher rates of attrition than their rural counterparts. For this and other reasons, most of the analysis is limited to the non-Nairobi sample.[24] Although this concern has merit, less is known about the impact of M-PESA on the risk-sharing capability of rural households in particular.

Table 17 reports summary statistics by rural–urban status over the two periods. These two groups are quite different in terms of variables such as demographics, financial services use and occupations. Most specifically, rural households' access to formal financial services, such as bank accounts, is relatively low. Interestingly, Table 17 shows there was a small decrease in the use of bank accounts for rural households between the two periods (grey-highlighted rows). Given the heterogeneity observed in the availability of alternative financial services in rural and urban area, it would be of great interest to see whether the risk-sharing impact of M-PESA adoption differs depending on the residence location.

There is relatively little literature that focuses on the use of financial institutions in rural areas. One exception is Munyegera and Matsumoto (2016). Their study was motivated by the belief that most formal financial institutions are concentrated in the urban areas, where people are better able to smooth their consumption against negative shocks. They find that remittances receipt is a crucial source of informal insurance for poor rural households in Uganda.

---

[23] Due to the limited coverage of cell phone towers and M-PESA agents, the residents of the north and northeast parts of the country were excluded from the sample.
[24] Nairobi is the capital and largest city of Kenya.

Table 17: Summary statistics (period 1, rural versus urban)

	Period 1 (rural)		Period 1 (urban)	
	Mean	SD	Mean	SD
M-PESA user (per cent)	0.289	0.453	0.532	0.499
Own cell phone (per cent)	0.552	0.498	0.789	0.408
Per capita consumption (Kenyan shillings)	41,530	42,620	94,594	163,219
Per capita food consumption (Kenyan shillings)	24,257	20,398	37,046	35,861
Total wealth (Kenyan shillings)	96,387	305,751	152,401	486,504
HH size	4.825	2.313	3.915	2.082
Education of head (years)	5.638	5.364	7.884	5.692
Positive shock (per cent)	0.100	0.301	0.115	0.319
Negative shock (per cent)	0.537	0.499	0.474	0.499
Weather/Agricultural shock (per cent)	0.065	0.247	0.019	0.135
Illness shock (per cent)	0.240	0.427	0.246	0.431
Send remittances (per cent)	0.346	0.476	0.543	0.498
Receive remittances (per cent)	0.363	0.481	0.403	0.491
Financial access dummies (per cent)				
Bank account	0.368	0.483	0.597	0.491
Mattress	0.773	0.419	0.750	0.433
Savings and credit cooperative (SACCO)	0.188	0.391	0.188	0.391
Rotating savings and credit cooperative (ROSCA)	0.411	0.492	0.400	0.490
Household head occupation dummies (per cent)				
Farmers	0.560	0.497	0.101	0.302
Public service	0.028	0.164	0.042	0.201
Professional occupation	0.128	0.335	0.304	0.460
Househelp	0.042	0.200	0.128	0.334
Run a business	0.090	0.286	0.184	0.388
Sales	0.020	0.141	0.068	0.252
In industry	0.033	0.178	0.032	0.175
Other occupation	0.037	0.188	0.076	0.265
Unemployed	0.059	0.237	0.065	0.246
Observations	691		1,591	

Table 17: Summary statistics (period 2, rural versus urban) (continued)

	Period 2 (rural)		Period 2 (urban)	
	Mean	SD	Mean	SD
M-PESA user (per cent)	0.599	0.490	0.766	0.423
Own cell phone (per cent)	0.675	0.467	0.815	0.388
Per capita consumption (Kenyan shillings)	37,070	35,420	82,708	105,510
Per capita food consumption (Kenyan shillings)	22,772	17,699	35,150	28,844
Total wealth (Kenyan shillings)	80,152	134,786	175,333	902,523
HH size	4.933	2.342	4.027	2.241
Education of head (years)	6.622	4.583	8.173	5.190
Positive shock (per cent)	0.074	0.262	0.061	0.239
Negative shock (per cent)	0.559	0.497	0.579	0.494
Weather/Agricultural shock (per cent)	0.170	0.376	0.109	0.312
Illness shock (per cent)	0.379	0.486	0.422	0.494
Send remittances (per cent)	0.336	0.473	0.551	0.498
Receive remittances (per cent)	0.419	0.494	0.420	0.494
Financial access dummies (per cent)				
Bank account	0.358	0.480	0.622	0.485
Mattress	0.804	0.398	0.712	0.453
Savings and credit cooperative (SACCO)	0.178	0.382	0.175	0.380
Rotating savings and credit cooperative (ROSCA)	0.461	0.499	0.460	0.499
Household head occupation dummies (per cent)				
Farmers	0.519	0.500	0.103	0.303
Public service	0.013	0.114	0.048	0.213
Professional occupation	0.130	0.336	0.241	0.428
Househelp	0.072	0.258	0.124	0.330
Run a business	0.116	0.320	0.194	0.396
Sales	0.049	0.215	0.121	0.326
In industry	0.007	0.086	0.027	0.161
Other occupation	0.025	0.158	0.055	0.227
Unemployed	0.069	0.253	0.083	0.275
Observations	691		1,591	

In order to see whether the impact of M-PESA adoption on consumption smoothing differed among the households who reside in rural areas compared to their urban counterparts, I focused the analysis on an exclusively rural sample. Table 18 presents results of basic specifications for three different samples, including full sample and rural and urban sub-samples. According to the baseline results, regardless of being an M-PESA user or nonuser, rural households experience similar reductions in per capita consumption (see the bottom rows). However, these two coefficients are not significantly different from zero.

The coefficient on the interaction of interest is no longer significant in panels 2 and 3, once I look at the rural and urban sub-samples. However, once a full set of control variables is considered (panels 4 and 5), users in urban areas are able to smooth shocks perfectly, and even experience an 8 per cent increase in consumption. Nonusers for the same sub-sample, on the other hand, suffer a 6 per cent (non-significant) reduction. Looking at the rural sub-sample when a full set of control variables is included, I find that users are able to smooth consumption, and experience a 3 per cent increase in consumption (far lower than the urban counterparts), whereas nonusers suffer an 8 per cent reduction (none of these estimates is significant).

Table 18: Basic difference-in-differences results, heterogeneous effects

	Total consumption					
	OLS (1)			Panel (2)		
	Original (1)	Rural (2)	Urban (3)	Original (4)	Rural (5)	Urban (6)
M-PESA user						
	0.5730***	0.4280***	0.4419***	0.0520	0.1275	0.0132
	[0.0377]	[0.0599]	[0.0467]	[0.0481]	[0.0840]	[0.0630]
Negative shock	–0.2111***	–0.0718***	–0.3509***	–0.0668	–0.0718	–0.0500
	[0.0381]	[0.0527]	[0.0511]	[0.0491]	[0.0720]	[0.0697]
User × shock	0.0917*	–0.0006	0.2202***	0.1093*	0.0857	0.0958
	[0.0506]	[0.0790]	[0.0634]	[0.0616]	[0.1022]	[0.0887]
Demographic controls						
Controls + interaction				Yes		
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Time × location FE						
Observations	4,562	1,382	3,180	4,562	1,382	3,180
Negative shock	–0.1593***	–0.0721*	–0.2079***	–0.0050	–0.0338	0.0122
	[0.0252]	[0.0393]	[0.0304]	[0.0305]	[0.0473]	[0.0419]
Shock, users	–0.1194***	–0.0725	–0.1308***	0.0425	0.0139	0.0457
	[0.0335]	[0.0590]	[0.0378]	[0.0379]	[0.0659]	[0.0533]
Shock, nonusers	–0.2111***	–0.0718	–0.3509***	–0.0668	–0.0717	–0.0500
	[0.0381]	[0.0527]	[0.0511]	[0.0491]	[0.0720]	[0.0697]
Shock, nonusers | user Xs						
Mean of user	0.5656	0.4440	0.6498	0.5656	0.4440	0.6498

Notes: Dependent variable is log total household consumption per capita. Heteroscedasticity-robust standard errors are reported in brackets. Controls: household demographics; household education and occupation; use of bank accounts, SACCOs and ROSCAs; cell phone ownership. Interactions refer to interactions of the controls with the shock. When interactions are included, the overall effect of shock is evaluated at the mean of the covariates. ***, **, * Significant at 1, 5 and 10 per cent levels.

The last row reports the effect for nonusers evaluated at mean characteristics of users. These two estimates are statistically different from zero.

Table 18: Basic difference-in-differences results, heterogeneous effects (continued)

	Total consumption						
	Panel(3)			Panel(4 and 5)			
	Original (7)	Rural (8)	Urban (9)	Original (10)	Original (11)	Rural (12)	Urban (13)
M-PESA user							
	0.0456	0.1327	0.0131	–0.0223	–0.0088	0.0776	–0.0751
	[0.0469]	[0.0811]	[0.0637]	[0.0484]	[0.0449]	[0.0807]	[0.0587]
Negative shock	–0.0727	–0.0842	–0.0552	0.2872	0.2673	0.0976	0.5581**
	[0.0468]	[0.0697]	[0.0654]	[0.1762]	[0.1799]	[0.2209]	[0.2296]
User × shock	0.1320**	0.1081	0.1239	0.1749***	0.1483**	0.1707	0.1989**
	[0.0594]	[0.1021]	[0.0848]	[0.0663]	[0.0599]	[0.1075]	[0.0784]
Demographic controls				Yes	Yes	Yes	Yes
Controls + interaction				Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time × location FE	Yes				Yes		
Observations	4,562	1,382	3,180	4,545	4,545	1,377	3,168
Negative shock	0.0019	–0.0362	0.0252	0.0022	–0.0059	–0.0320	0.0312
	[0.0292]	[0.0463]	[0.0407]	[0.0286]	[0.0280]	[0.0436]	[0.0356]
Shock, users	0.0592	0.0239	0.0686	0.0518	0.0460	0.0296	0.0804*
	[0.0370]	[0.0674]	[0.0526]	[0.0383]	[0.0355]	[0.0645]	[0.0439]
Shock, nonusers	–0.0727	0.0842	–0.0552	–0.0626	–0.0737*	–0.0813	–0.0604
	[0.0468]	[0.0697]	[0.0654]	[0.0447]	[0.0429]	[0.0629]	[0.0574]
Shock, nonusers | user Xs				–0.1230**	–0.1024**	–0.1411*	–0.1186*
				[0.0549]	[0.0502]	[0.0838]	[0.0669]
Mean of user	0.5656	0.4440	0.6498	0.5661	0.5656	0.4441	0.6506

Notes: Columns 10 and 11 are columns 4 and 5 in Table 4A. Given that I run separate estimations for urban and rural sample, I exclude the location-level control variables. As a result, panels 4 and 5 become identical. ***, **, * Significant at 1, 5 and 10 per cent levels.

Given that households who reside within a village may experience village-level aggregate shocks, I check the robustness of the research findings by clustering at the village level. The results are reported in Table 19. Looking at the last column, which includes the full set of controls, I find that the effects for users and nonusers are no longer significantly different from zero.

Therefore, the theory of change analysis does not support the idea that the risk-sharing impact of M-PESA adoption differs depending on the residence location.

Table 19: Basic difference-in-differences results, multilevel fixed effects with multiway clustering (clustered at households and village levels)

	Total consumption Full sample							
	Panel (2)		Panel (3)		Panel (4)		Panel (5)	
	Original panel (1)	MFE-MC panel (2)	Original panel (3)	MFE-MC panel (4)	Original panel (5)	MFE-MC panel (6)	Original panel (7)	MFE-MC panel (6)
M-PESA user	0.0520	0.0621	0.0456	0.0572	–0.0223	–0.0222	–0.0088	0.0088
	[0.0481]	[0.0531]	[0.0469]	[0.0525]	[0.0484]	[0.0543]	[0.0449]	[0.0523]
Negative shock	–0.0688	–0.0512	–0.0727	–0.0583	0.2872	0.2872	0.2673	0.2895
	[0.0491]	[0.0543]	[0.0468]	[0.0528]	[0.1762]	[0.1662]	[0.1799]	[0.1857]
User × shock	0.1093*	0.0922	0.1320**	0.1174*	0.1749***	0.1749***	0.1483**	0.1305**
	[0.0616]	[0.0604]	[0.0594]	[0.0728]	[0.0663]	[0.0646]	[0.0599]	[0.0633]
Demographic controls			Yes	Yes	Yes	Yes	Yes	Yes
Controls + Interactions					Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time × District FE	Yes	Yes	Yes	Yes			Yes	Yes
Observations	4,562	4,362	4,562	4,362	4,545	4,526	4,545	4,328
Negative Shock	–0.0050	0.0007	0.0019	0.0077	0.0022	0.0022	–0.0059	–0.0033
	[0.0305]	[0.0348]	[0.0292]	[0.0332]	[0.0286]	[0.0306]	[0.0280]	[0.0321]
Shock, users	0.0425	0.0410	0.0592	0.0591	0.0518	0.0515	0.0460	0.0441
	[0.0379]	[0.0382]	[0.0370]	[0.0375]	[0.0383]	[0.0396]	[0.0355]	[0.0358]
Shock, nonusers	–0.0668	–0.0512	–0.0727	–0.0583	–0.0626	–0.0623	–0.0737*	–0.0644
	[0.0491]	[0.0543]	[0.0468]	[0.0528]	[0.0447]	[0.0478]	[0.0429]	[0.0500]
Shock, nonusers | user Xs					–0.1230**	–0.1233**	–0.1024**	–0.0864
					[0.0549]	[0.0575]	[0.0502]	[0.0562]
Mean of user	0.5656	0.5626	0.5656	0.5626	0.5661	0.5666	0.5661	0.5633

Notes: MFE-MC refers to multilevel fixed effects, multiway clustering. Two levels of clustering include households and locations (reghdfe command is used).
***, **, * Significant at 1, 5 and 10 per cent levels.

5. Conclusions

Jack and Suri (2014) explore the impact of mobile money on informal risk sharing. They find that the introduction of M-PESA facilitates the redistribution of finances across geographical distances. This has improved households' ability to spread risk, which is mainly attributed to the associated reduction in transaction costs. As a result, M-PESA has been able to lift many Kenyan households out of poverty (Suri and Jack 2016).

My replication study consisted of three parts: (i) push-button and pure replication, (ii) measurement and estimation analysis and (iii) theory of change analysis. First, I employed the procedures of push-button replication and the 'independent coding' method of pure replication to study Jack and Suri (2014). I was able to reproduce all of the main findings from the original paper, although I discovered some minor differences in the code and tables. Accordingly, my pure and push-button replication is categorised as comparable but incomplete.

Second, I examined whether their results were robust to a number of changes. I first looked at the data and checked the existence of outliers. I then assessed the robustness of findings to alternative statistical and estimation methods. I interpret my results as indicating that the original results are robust to different specification and approaches.

Finally, given the heterogeneity observed in the availability of alternative financial services in rural and urban areas, it is of great interest to see whether the risk-sharing impact of M-PESA adoption differed depending on the residence location. Therefore, I changed the focus of analysis from the entire population, including rural and urban residents, to the rural segment of the population. The latter is of interest because this population is supposedly excluded by formal financial services. This examination addresses M-PESA promise of providing 'banking for the unbanked'. However, my analysis did not identify significant differences between urban and rural residents in the benefits received from M-PESA.

Jack and Suri's findings and my replication provide strong empirical evidence that M-PESA has had a positive impact on people's financial health in Kenya. The financial benefits derived from such mobile money innovations can play a vital role in combatting world poverty. Policymakers should take one crucial step – to adopt enabling policies that allow competing firms to offer innovative new technologies such as M-PESA.

Appendix A: Tables

Table A1: Replication results of summary statistics (full sample), a reproduction of Jack and Suri's Table 1A

	Round1		Round2	
	Mean	SD	Mean	SD
M-PESA user (per cent)	0.433	0.496	0.698	0.459
Own cell phone (per cent)	0.692	0.462	0.758	0.428
Per capita consumption (Kenyan shillings)	72,883	131,000	64,017	87,115
Per capita food consumption (Kenyan shillings)	31,814	31,134	30,081	25,621
Total wealth (Kenyan shillings)	129,482	422,829	136,377	700,497
HH size	4.287	2.224	4.398	2.325
Education of head (years)	6.967	5.668	7.537	5.007
Positive shock (per cent)	0.109	0.312	0.066	0.249
Negative shock (per cent)	0.500	0.500	0.571	0.495
Weather / agricultural shock (per cent)	0.038	0.190	0.134	0.341
Illness shock (per cent)	0.243	0.429	0.404	0.491
Send remittances (per cent)	0.463	0.499	0.463	0.499
Receive remittances (per cent)	0.387	0.487	0.420	0.494
Financial access dummies (per cent)				
Bank account	0.504	0.500	0.514	0.500
Mattress	0.759	0.428	0.750	0.433
Savings and credit cooperative (SACCO)	0.188	0.391	0.176	0.381
Rotating savings and credit cooperative (ROSCA)	0.404	0.491	0.460	0.498
Household head occupation dummies (per cent)				
Farmers	0.289	0.453	0.273	0.446
Public service	0.036	0.187	0.034	0.180
Professional occupation	0.232	0.422	0.196	0.397
Househelp	0.093	0.290	0.103	0.304
Run a business	0.146	0.353	0.162	0.369
Sales	0.049	0.215	0.091	0.288
In industry	0.032	0.176	0.019	0.136
Other occupation	0.060	0.237	0.043	0.202
Unemployed	0.062	0.242	0.077	0.266
Observations	2,282		2,282	

Table A2: Replication results of remittances for non-Nairobi sample (only means reported), a reproduction of Jack and Suri's Table 1B

	Round1		Round2	
	Sent	Received	Sent	Received
Overall remittances				
Number of remittances per month	1.303	0.862	1.075	0.815
Total value	4,531	5,025	3,136	2,134
Total value (fraction of consumption)	0.016	0.019	0.015	0.012
Average distance (Km)	234.1	288.4	213.7	235.0
Net value remitted	1,494.3		–483.8	
M-PESA remittances				
Remittances	0.933	0.807	1.615	0.847
Total value	7,965.4	9,923.7	7,711.3	4,789.7
Average distance (Km)	343.6	335.1	238.2	237.3
Non-M-PESA remittances				
Remittances	1,930	1.402	0.760	1.080
Total value	9,717.3	13,694.3	4,614.5	5,057.5
Average distance (km)	194.2	273.3	172.4	230.8

Table A3: Replication results of remittances for non-Nairobi sample (only means reported), a reproduction of Jack and Suri's Table 1B (updated version)

	Round1		Round2	
	Sent	Received	Sent	Received
Overall remittances				
Number of remittances per month	2.860	2.209	2.375	1.929
Total value	10,073	13,019	6,947	5,094
Total value (fraction of consumption)	0.036	0.050	0.032	0.029
Average distance (km)	234.1	288.4	213.7	235.0
Net value remitted	2,355.9		–789.8	
M-PESA remittances				
Remittances	0.933	0.807	1.615	0.847
Total value	7,965.4	9,923.7	7,711.3	4,789.7
Average distance (km)	343.6	335.1	238.2	237.3
Non-M-PESA remittances				
Remittances	1,930	1.402	0.760	1.080
Total value	9,717.3	13,694.3	4,614.5	5,057.5
Average distance (km)	194.2	273.3	172.4	230.8

Table A4: Replication results of remittances received for non-Nairobi sample, a reproduction of Jack and Suri's Table 1C

Method money/transfer was sent	Frequency (per cent)	Average cost
Hand delivery by self	14.8	1.69
Hand delivery by friend	4.8	2.51
Bus delivery through friend/relative	5.3	8.85
Bus delivery through driver/courier	3.5	144.85
Western Union	0.7	99.29
M-PESA from own/friend's/agent's account	59.3	51.35
Postal bank	3.9	184.30
Direct deposit	4.8	104.78
Other	2.8	69.30

Table A5: Replication results of agent characteristics, a reproduction of Jack and Suri's Table 2

	Full sample				Non-Nairobi Sample			
	Round1		Round2		Round1		Round2	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Panel A. Household access to agents								
Number agents w/in 1 km	3.31	7.15	6.99	15.06	2.55	5.28	5.08	10.07
Number agents w/in 2 km	9.37	29.10	19.60	58.86	4.63	8.03	9.78	17.34
Number agents w/in 5 km	29.67	92.49	60.18	178.0	9.71	19.09	21.81	47.35
Number agents w/in 10 km	60.94	173.2	127.8	344.7	18.71	43.32	45.11	103.8
Number agents w/in 20 km	115.3	275.1	240.7	544.7	54.11	150.01	120.5	301.9
Dist to closest agent (km)	4.86	7.96	3.98	7.25	5.04	7.49	4.13	6.87
Log dist to closest agent	7.36	1.65	7.11	1.65	7.47	1.61	7.23	1.61

Panel B. Agent distribution

	Full sample	
	Round 1	Round 2
Difference in distance between closest and second closest agent (per cent of distance to closest agent)		

Panel C. Agent-level data (total number of agents=7,685)

Agent business	Mean	SD
New registration, past 7 days	7.012	8.782
Transactions, past 7 days	70.687	49.357
Frequency of stockouts	E-money stockout (per cent)	Cash stockout (per cent)
At least one every 2 weeks	30.82	15.99
Once a month	8.50	4.47
Less often than that	3.44	3.48
Never	57.24	76.06

Notes: There is a minor difference for the total number of agents (light grey–highlighted area). The dark grey–highlighted cells show the cases I am unable to replicate due to data unavailability.

Table A6: Replication results of correlates of shock measures, a reproduction of Jack and Suri's Table 3

	Overall shock			Illness shock		
	Coefficient	SE	Partial R^2	Coefficient	SE	Partial R^2
M-PESA user	–0.0117	[0.0380]	0.0000	0.0078	[0.0344]	0.0005
Cell phone ownership	–0.0127	[0.0437]	0.0035	0.0008	[0.0382]	0.0014
Log distance to agent	0.0061	[0.0516]	0.0002	–0.0584	[0.0511]	0.0000
Agents within 1 km	–0.0470	[0.0504]	0.0007	–0.0123	[0.0483]	0.0010
Agents within 2 km	0.0774	[0.0544]	0.0000	–0.0377	[0.0524]	0.0002
Agents within 5 km	0.0309	[0.0375]	0.0001	0.0618*	[0.0319]	0.0000
Occupation-farmer	0.0520	[0.0608]	0.0028	0.0310	[0.0592]	0.0014
Occupation-professional	0.0519	[0.0580]	0.0002	–0.0006	[0.0555]	0.0000
Occupation-househelp	0.0262	[0.0653]	0.0001	0.0146	[0.0633]	0.0003
Occupation-run a Business	–0.0167	[0.0615]	0.0000	–0.0766	[0.0583]	0.0000
Occupation-sales	0.0075	[0.0710]	0.0005	–0.0805	[0.0615]	0.0013
Occupation-unemployed	0.1050	[0.0741]	0.0021	0.0512	[0.0670]	0.0010
HH has a bank account	0.0155	[0.0382]	0.0001	0.0172	[0.0346]	0.0000
HH has a ROSCA account	–0.0078	[0.0310]	0.0014	0.0047	[0.0277]	0.0022
HH has a SACCO account	0.0545	[0.0417]	0.0008	0.0051	[0.0346]	0.0000
Fraction of boys in HH	–0.0799	[0.1621]	0.0002	–0.1612	[0.1461]	0.0000
Fraction of girls in HH	0.0158	[0.1421]	0.0006	0.0289	[0.1234]	0.0008
HH size	0.0195	[0.0148]	0.0057	0.0032	[0.0147]	0.0018
F Statistics [*P*-value]	0.84 [0.6543]			0.87 [0.6142]		

Notes: The grey-highlighted cells indicate minor differences in the partial R^2 produced by Jack and Suri's code and the values reported in their table.

Table A7: Replication results of basic difference-in-differences results, a reproduction of Jack and Suri's Table 4A

	Total consumption Full sample				
	OLS	Panel	Panel	Panel	Panel
	(1)	(2)	(3)	(4)	(5)
M-PESA user	0.5730***	0.0520	0.0456	–0.0223	–0.0088
	[0.0377]	[0.0481]	[0.0469]	[0.0484]	[0.0449]
Negative shock	–0.2111***	–0.0668	–0.0727	0.2872	0.2673
	[0.0381]	[0.0491]	[0.0468]	[0.1762]	[0.1799]
User × negative shock	0.0917*	0.1093*	0.1320**	0.1749***	0.1483**
	[0.0506]	[0.0616]	[0.0594]	[0.0663]	[0.0599]
Demographic controls			Yes	Yes	Yes
Controls + interactions				Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Time × location FE		Yes	Yes		Yes
Observations	4,562	4,562	4,562	4,545	4,545
Negative shock	–0.1593***	–0.0050	0.0019	0.0022	–0.0059
	[0.0252]	[0.0305]	[0.0292]	[0.0286]	[0.0280]
Shock, users	–0.1194***	0.0425	0.0592	0.0518	0.0460
	[0.0335]	[0.0379]	[0.0370]	[0.0383]	[0.0355]
Shock, nonusers	–0.2111***	–0.0668	–0.0727	–0.0626	–0.0737*
	[0.0381]	[0.0491]	[0.0468]	[0.0447]	[0.0429]
Shock, nonusers | user Xs				–0.1230**	–0.1024**
				[0.0549]	[0.0502]
Mean of user	0.5656	0.5656	0.5656	0.5661	0.5661

Table A8: Replication results of basic difference-in-differences results, a reproduction of Jack and Suri's Table 4B

	Total consumption panel without Nbi (1)	Total consumption panel without Nbi (2)	Food consumption panel without Nbi (3)	Total consumption panel without Msa (4)	Total consumption panel without Msa (4-new)	Total consumption panel poor (5)
M-PESA user	–0.0161	0.0020	0.0174	0.0231	0.0105	–0.0564
	[0.0511]	[0.0470]	[0.0431]	[0.0489]	[0.0466]	[0.0546]
Negative shock	0.1865	0.1544	0.0749	0.1458	0.2665	0.2711
	[0.1502]	[0.1627]	[0.1389]	[0.1697]	[0.1877]	[0.2110]
User × negative shock	0.1784**	0.1380**	0.0586	0.1404**	0.1508**	0.2068***
	[0.0700]	[0.0632]	[0.0636]	[0.0654]	[0.0616]	[0.0764]
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Controls + interactions	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Time × location FE		Yes	Yes	Yes	Yes	Yes
Observations	3,911	3,911	3,908	3,703	4,337	2,723
Negative shock	0.0045	–0.0041	–0.0335	–0.0065	–0.0072	0.0206
	[0.0301]	[0.0294]	[0.0275]	[0.0302]	[0.0288]	[0.0351]
Shock, users	0.0516	0.0415	–0.0124	0.0399	0.0460	0.1273***
	[0.0409]	[0.0375]	[0.0331]	[0.0385]	[0.0363]	[0.0458]
Shock, nonusers	–0.0533	–0.0601	–0.0594	–0.0626	–0.0758	–0.0755
	[0.0459]	[0.0442]	[0.0435]	[0.0456]	[0.0441]	[0.0520]
Shock, nonusers | user Xs	–0.1267**	–0.0965*	–0.0710	–0.1005*	–0.1049*	–0.0795
	[0.0585]	[0.0532]	[0.0551]	[0.0554]	[0.0520]	[0.0611]
Mean of user	0.5512	0.5512	0.5514	0.5470	0.5629	0.4739

Notes: The sample corresponding to column 4 excludes both Mombasa and Nairobi. Column 4–new, re-estimated the specification of column 4, excluding Mombasa, but keeping Nairobi.

Table A9: Replication results of health shocks (panel), a reproduction of Jack and Suri's Table 4C

	Total consumption Illness shock			Nonhealth consumption Illness shock		
	(1)	(2)	(3)	(4)	(5)	(6)
M-PESA user	0.0830**	0.0298	0.0493	0.0885**	0.0366	0.0577
	[0.0417]	[0.0410]	[0.0411]	[0.0402]	[0.0395]	[0.0396]
Negative shock	–0.0135	0.1244	0.1652	–0.0826	–0.0160	–0.0032
	[0.0521]	[0.1845]	[0.1690]	[0.0507]	[0.1822]	[0.1682]
User × negative shock	0.1354**	0.1691**	0.0921	0.1531**	0.1757**	0.1057
	[0.0654]	[0.0686]	[0.06871]	[0.0634]	[0.0645]	[0.0652]
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Controls + interactions		Yes	Yes		Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Time × rural FE	Yes	Yes	Yes	Yes	Yes	Yes
Time × location FE			Yes			Yes
Observations	4,562	4,545	4,545	4,562	4,545	4,545
R^2	0.083	0.148	0.312	0.090	0.155	0.318
Shock effect	0.0631*	0.0481	0.0415	0.0041	–0.0088	–0.0155
	[0.0326]	[0.0319]	[0.0305]	[0.0314]	[0.0306]	[0.0297]
Shock, users	0.1220***	0.1157***	0.0939*	0.0705	0.0667	0.0463
	[0.0406]	[0.0403]	[0.0386]	[0.0389]	[0.0389]	[0.0377]
Shock, nonusers	–0.0135	–0.0400	–0.0268	–0.0826	–0.1075**	–0.0961**
	[0.0521]	[0.0488]	[0.0461]	[0.0507]	[0.0469]	[0.0444]
Shock, nonusers | user Xs		–0.0534	0.0018		–0.1089*	–0.0594
		[0.0577]	[0.0591]		[0.0536]	[0.0560]
Mean of shock	0.3240	0.3240	0.3240	0.3240	0.3240	0.3240

Notes: Given that the methodology chosen for pure replication is independent coding, the whole sample for these analyses is considered. However, closer investigation uncovered that the results in the source table excluded Nairobi.

Table A10: Replication results of health shocks (panel), a reproduction of Jack and Suri's Table 4C (updated version)

	Total consumption Illness shock			Nonhealth consumption Illness shock		
	(1)	(2)	(3)	(4)	(5)	(6)
M-PESA user	0.0978**	0.0386	0.0618	0.1037**	0.0459	0.0688*
	[0.0438]	[0.0434]	[0.0434]	[0.0422]	[0.0419]	[0.0417]
Negative shock	–0.0045	–0.0260	–0.0104	–0.0759	–0.1643	–0.1754
	[0.0527]	[0.1589]	[0.1515]	[0.0514]	[0.1627]	[0.1550]
User × negative shock	0.1190*	0.1585**	0.0630	0.1380**	0.1641**	0.0780
	[0.0671]	[0.0728]	[0.0731]	[0.0651]	[0.0684]	[0.0694]
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Controls + interactions		Yes	Yes		Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Time × rural FE	Yes	Yes	Yes	Yes	Yes	Yes
Time × location FE			Yes			Yes
Observations	3,927	3,911	3,911	3,927	3,911	3,911
R^2	0.088	0.150	0.323	0.096	0.157	0.329
Shock effect	0.0610*	0.0466	0.0367	0.0001	–0.0121	–0.0228
	[0.0333]	[0.0331]	[0.0320]	[0.0323]	[0.0318]	[0.0311]
Shock, users	0.1145***	0.1104***	0.0781*	0.0621	0.0604	0.0293
	[0.0421]	[0.0423]	[0.0406]	[0.0406]	[0.0409]	[0.0398]
Shock, nonusers	–0.0045	–0.0316	–0.0142	–0.0759	–0.1011**	–0.0868*
	[0.0527]	[0.0503]	[0.0477]	[0.0514]	[0.0483]	[0.0460]
Shock, nonusers | user Xs		–0.0482	0.0152		–0.1037*	–0.0488
		[0.0611]	[0.0629]		[0.0567]	[0.0594]
Mean of shock	0.3231	0.3231	0.3231	0.3231	0.3231	0.3231

Notes: This table repeats the estimation reported in previous table (Table A9) but for the non-Nairobi sample instead.

Table A11: Replication results of mechanisms (panel), a reproduction of Jack and Suri's Table 5A

	Overall shock: sample w /out Nairobi				Overall shock: w /out Mombasa		Illness shock	
	Pr [receive]		Number received	Total received (square root)	Pr[receive]	Total received (square root)	Pr[receive]	Total received (square root)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
M-PESA user	0.1897***	0.1528***	0.2574**	10.6757***	0.1143**	9.0579**	0.1726***	12.5548***
	[0.0456]	[0.0487]	[0.1305]	[3.7863]	[0.0517]	[4.0683]	[0.0420]	[3.1596]
Negative shock	–0.0442	–0.0409	–0.1306	1.8775	–0.1027	–1.8885	–0.1417	–9.3597
	[0.0390]	[0.1438]	[0.4193]	[12.0864]	[0.1452]	[12.4371]	[0.1457]	[10.9683]
User × shock	0.0923*	0.1337**	0.3286*	8.3428*	0.1733***	10.0472**	0.1598**	8.6003
	[0.0530]	[0.0633]	[0.1789]	[4.6884]	[0.0666]	[4.9200]	[0.0722]	[5.2788]
Controls + interactions		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time × location FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,928	3,911	3,911	3,873	3,703	3,665	3,911	3,873
R^2	0.199	0.218	0.184	0.203	0.223	0.205	0.223	0.209
Shock effect	0.0066	0.0099	–0.0369	1.6647	0.0043	1.5026	0.0161	2.7412
	[0.0282]	[0.0288]	[0.0871]	[2.2697]	[0.0297]	[2.3569]	[0.0315]	[2.5233]
Shock, users	0.0481	0.0478	0.0470	4.3755	0.0543	4.6901	0.0735*	6.5410*
	[0.0383]	[0.0381]	[0.1157]	[3.4195]	[0.0391]	[3.5678]	[0.0433]	[3.5215]
Shock, nonusers	–0.0442	–0.0366	–0.1400	–1.6403	–0.0561	–2.3154	–0.0544	–1.8914
	[0.0390]	[0.0407]	[0.1221]	[2.6656]	[0.0425]	[2.7528]	[0.0442]	[3.0544]
Mean of user	0.5504	0.5512	0.5512	0.5494	0.5470	0.5450	0.5512	0.5494

Table A12: Replication results of where remittances come from: distance and the role of networks (panel), a reproduction of Jack and Suri's Table 5B

	Log distance travelled		Number of different senders		Fraction of network remitting	
	Overall shock	Illness shock	Overall shock	Illness shock	Overall shock	Illness shock
	(1)	(2)	(3)	(4)	(5)	(6)
M-PESA user	0.0460	–0.0980	0.1783***	0.2004***	0.1012***	0.1128***
	[0.4424]	[0.3435]	[0.0678]	[0.0551]	[0.0363]	[0.0319]
Shock	–0.2546	–0.2050	–0.3071	–0.4348*	–0.0675	–0.1490
	[0.7437]	[0.9028]	[0.2160]	[0.2227]	[0.1279]	[0.1252]
User × shock	0.2279	1.3929**	0.2008**	0.2519***	0.0936*	0.1090*
	[0.5653]	[0.6446]	[0.0874]	[0.0968]	[0.0493]	[0.0612]
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Controls + interactions	Yes	Yes	Yes	Yes	Yes	Yes
Time × location FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,518	1,518	3,911	3,911	3,394	3,394
R^2	0.484	0.488	0.194	0.200	0.241	0.246
Shock effect	–0.3303	–0.2001	0.0249	0.0270	0.0125	0.0246
	[0.2166]	[0.2355]	[0.0439]	[0.0457]	[0.0221]	[0.0236]
Shock, users	–0.3076	0.0890	0.0852	0.1123*	0.0325	0.0484*
	[0.2563]	[0.2764]	[0.0580]	[0.0606]	[0.0247]	[0.0270]
Shock, nonusers	–0.4026	–1.1204**	–0.0493	–0.0779	–0.0188	–0.0128
	[0.4150]	[0.5099]	[0.0594]	[0.0620]	[0.0384]	[0.0441]
Mean of user	0.7609	0.7609	0.5512	0.5512	0.6104	0.6104

Table A13: Replication results of reduced forms using agent rollout (panel), a reproduction of Jack and Suri's Table 6A

	Overall shock Agents w / in 1km			Illness shock Agents w / in 1km
	(1)	(2)	(3)	(4)
Negative shock	–0.0525	–0.0543	–0.0543	–0.0591
	[0.0470]	[0.0464]	[0.0410]	[0.0425]
Agents	–0.0331	–0.0210	0.0450	0.0552
	[0.0400]	[0.0382]	[0.0377]	[0.0381]
Agents × shock	0.0470**	0.0534***	0.0451**	0.0350
	[0.0220]	[0.0199]	[0.0177]	[0.0216]
Controls		Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Time × Location FE			Yes	Yes
Time × Rural FE	Yes	Yes	Yes	Yes
Observations	3,927	3,911	3,911	3,911
R^2	0.015	0.135	0.305	0.312
Shock effect	0.0002	0.0055	–0.0037	–0.0199
	[0.0346]	[0.0332]	[0.0328]	[0.0346]
Mean of agents	1.1197	1.1206	1.1206	1.1206

Table A14: Replication results of reduced forms using agent rollout (panel), a reproduction of Jack and Suri's Table 6B

	Overall shock					
	Agents w / in 2 km		Agents w /in 5 km	Agents w / in 20 km	Distance to closest agent	
	(1)	(2)	(3)	(4)	(5)	(6)
Negative shock	–0.0706	–0.0739	–0.0405	–0.0154	0.3317**	0.3398***
	[0.0530]	[0.0460]	[0.0464]	[0.0559]	[0.1353]	[0.1294]
Agents	–0.0232	0.0010	–0.0021	–0.0130	–0.0096	0.0151
	[0.0352]	[0.0383]	[0.0258]	[0.0174]	[0.0438]	[0.0505]
Agents × Shock	0.0414**	0.0402***	0.0130	0.0014	–0.0450**	–0.0466***
	[0.0173]	[0.0144]	[0.0106]	[0.0069]	[0.0192]	[0.0174]
Controls		Yes	Yes	Yes		Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Time × Location FE		Yes	Yes	Yes		Yes
Time × Rural FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,927	3,911	3,911	3,911	3,927	3,911
R^2	0.016	0.305	0.301	0.300	0.016	0.304
Shock effect	0.0023	–0.0031	–0.0046	–0.0058	0.0012	–0.0026
	[0.0345]	[0.0329]	[0.0334]	[0.0340]	[0.0344]	[0.0334]
Mean of agents	1.7613	1.7603	2.7539	6.7197	7.3486	7.3499

Table A15: Replication results of agent rollout, a reproduction of Jack and Suri's Table 6C

	Agents w/in 1 km		Agents w/in 2 km		Agents w/in 5 km		Distance to agent	
	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE
Log wealth	0.0047	[0.0088]	0.0155*	[0.0093]	–0.0079	[0.0123]	0.0079	[0.0061]
Cell phone ownership	–0.0288*	[0.0175]	0.0074	[0.0232]	–0.0183	[0.0286]	0.0040	[0.0180]
Household size	–0.0054	[0.0067]	0.0021	[0.0076]	–0.0073	[0.0105]	–0.0055	[0.0044]
Fraction of boys in the household	0.0559	[0.0794]	0.1005	[0.0987]	–0.0620	[0.1313]	0.0202	[0.0507]
Fraction of girls in the household	0.0868	[0.0700]	0.1226	[0.0847]	0.3236*	[0.1684]	–0.0286	[0.0613]
Occupation of head: farmer	0.0290	[0.0189]	–0.0253	[0.0216]	0.0211	[0.0233]	0.0044	[0.0157]
Occupation of head: professional	0.0082	[0.0304]	0.0420	[0.0391]	–0.0036	[0.0413]	0.0184	[0.0196]
Occupation of head: business	–0.0409	[0.0276]	0.0232	[0.0302]	0.0226	[0.0418]	–0.0009	[0.0200]
Households years of education	–0.0033	[0.0021]	–0.0008	[0.0026]	0.0040	[0.0031]	–0.0018	[0.0014]
HH has a bank account	0.0181	[0.0184]	0.0151	[0.0238]	0.0316	[0.0300]	0.0178	[0.0111]
HH has a SACCO account	0.0011	[0.0237]	–0.0061	[0.0276]	0.0327	[0.0505]	–0.0042	[0.0185]
HH has a ROSCA	0.0172	[0.0180]	0.0238	[0.0224]	0.0019	[0.0310]	0.0149	[0.0102]
Negative shock	0.0120	[0.0151]	0.0393**	[0.0183]	0.0492*	[0.0258]	–0.0035	[0.0120]
Illness shock	0.0004	[0.0171]	0.0008	[0.0205]	0.0433	[0.0256]	–0.0186	[0.0125]
	Agents w/in 1 km		**Agents w/in 2 km**		**Agents w/in 5 km**		**Agents w/in 1 km**	
	Period 1	Change	Period 1	Change	Period 1	Change	Period 1	Change
Distance to Nairobi	–0.0009 [0.0031]	0.0002 [0.0013]	–0.0026 [0.0058]	–0.0011 [0.0028]	–0.0029 [0.0099]	0.0028 [0.0047]	–0.0007 [0.0056]	–0.0003 [0.0011]

Notes: The grey-highlighted cells indicate cases in which the replicated results differ from the original estimates.

Table A16: Replication results of falsification test, 1997–2007, a reproduction of Jack and Suri's Table 7A

	Agents w/in 2 km				Distance to closest agent			
	Maize consumption		Crop consumption		Maize consumption		Crop consumption	
	OLS (1)	Panel (2)	OLS (3)	Panel (4)	OLS (5)	Panel (6)	OLS (7)	Panel (8)
Shock × agents	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Shock measure (positive measure)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Agents	[redacted]		[redacted]		[redacted]		[redacted]	
Observations	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
R^2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Table A17: Replication results of falsification test: similar sample for 2008–2009, a reproduction of Jack and Suri's Table 7B

	Using M-PESA user status		Using measures of agent access			
	Total consumption	Food consumption	Total consumption		Food consumption	
			Distance to agent	Agents w/in 2 km	Distance to agent	Agents w/in 2 km
	(1)	(2)	(3)	(4)	(5)	(6)
User/agent measure	0.0148	0.0012	0.0061	–0.0116	–0.0154	–0.0099
	[0.0628]	[0.0563]	[0.1074]	[0.0631]	[0.1108]	[0.0650]
Negative shock	0.1621	0.0191	0.7810***	0.0698	0.7184***	–0.1193
	[0.1670]	[0.1643]	[0.2690]	[0.1729]	[0.2506]	[0.1677]
User/Agent × Shock	0.1798**	0.0968	–0.0775***	0.0873***	–0.0911***	0.1045***
	[0.0803]	[0.0828]	[0.0291]	[0.0242]	[0.0260]	[0.0231]
Controls + Interactions	Yes	Yes	Yes	Yes	Yes	Yes
Time × Location FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,875	1,874	1,875	1,875	1,874	1,874
R^2	0.356	0.383	0.353	0.356	0.391	0.396
Negative shock effect	–0.0359	–0.0596	–0.0269	–0.0273	–0.0488	–0.0490
	[0.0411]	[0.0381]	[0.0415]	[0.0411]	[0.0378]	[0.0373]
Shock, users	0.0302	–0.0114				
	[0.0484]	[0.0456]				
Shock, nonusers	–0.0934*	–0.1014*				
	[0.0560]	[0.0539]				

Table A18: Replication results of correlates of nonattrition, a reproduction of Jack and Suri's Table 8A

	Full sample	Without Nairobi		Full sample	Without Nairobi
Log total expenditure	–0.0125	–0.0127	Years of education of	–0.0002	0.0012
	[0.0188]	[0.0199]	HH Head	[0.0019]	[0.0020]
M-PESA user	0.0337	0.0191	Occupation-farmer	0.0453	0.0438
	[0.0243]	[0.0260]		[0.0305]	[0.0314]
Cell phone	0.0467*	0.0544**	Occupation-	0.0263	0.0169
ownership	[0.0249]	[0.0257]	professional	[0.0306]	[0.0336]
Log distance to agent	–0.0025	–0.0112	Occupation-	–0.0150	–0.0170
	[0.0152]	[0.0177]	househelp	[0.0404]	[0.0446]
Agents within 1 km	–0.0092	–0.0074	Occupation-run a	0.0363	0.0309
	[0.0152]	[0.0165]	business	[0.0331]	[0.0350]
Agents within 2 km	–0.0155	–0.0322*	Occupation-sales	0.1001*	0.0860
	[0.0123]	[0.0181]		[0.0535]	[0.0563]
Agents within 5 km	0.0036	0.0036	HH has a bank	0.0231	0.0197
	[0.0080]	[0.0140]	account	[0.0226]	[0.0236]
Negative shock	0.0081	0.0140	HH has a ROSCA	0.0116	0.0131
	[0.0238]	[0.0255]	account	[0.0200]	[0.0210]
Illness shock	0.0091	0.0153	HH has a SACCO	0.0033	0.0104
	[0.0266]	[0.0281]		[0.0280]	[0.0295]
Sent remittance	0.0008	–0.0007	Household size	0.0141**	0.0143**
	[0.0213]	[0.0227]		[0.0055]	[0.0058]
Received remittance	–0.0192	–0.0163	Urban dummy	–0.0887**	–0.0791**
	[0.0217]	[0.0229]		[0.0366]	[0.0377]
Observations				2,998	2,518
R^2				0.168	0.176
F-Statistics [*p*-value]				2.47	2.58
				[0.0002]	[0.0001]

Notes: Significant difference between the original findings and replication results are highlighted grey in the table.

Table A19: Further results on attrition, a reproduction of Jack and Suri's Table 8B

	FGM weights			Limited sample where attrition is low at community level				
	M-PESA user		Dist to agent	M-PESA user			Agents w/in 1km	Dist to agent
	Total Consumption	Pr [receive]	Total consumption	Total consumption	Pr [receive]	Total received	Total consumption	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
User/agent measure	0.0036	0.1675***	0.0188	–0.0594	0.0994*	8.7203**	0.0739	0.0268
	[0.0472]	[0.0477]	[0.0534]	[0.0580]	[0.0581]	[4.2328]	[0.0566]	[0.0598]
Negative shock	0.1761	–0.0111	0.5503***	0.2015	–0.1031	–4.5028	0.1825	0.8997***
	[0.1639]	[0.1405]	[0.2100]	[0.1750]	[0.1967]	[13.0864]	[0.1793]	[0.2421]
User/agent × shock	0.1305**	0.1232**	–0.0499***	0.2469***	0.2380***	13.8458***	0.0843***	–0.0876***
	[0.0632]	[0.0627]	[0.0183]	[0.0740]	[0.0753]	[5.2629]	[0.0267]	[0.0242]
Controls + interactions	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time × location FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,893	3,893	3,893	2,789	2,789	2,761	2,789	2,789
R^2	0.326	0.224	0.326	0.357	0.240	0.228	0.357	0.357
Shock effect	–0.0034	0.0100	0.0004	–0.0007	0.0015	1.7150	0.0035	0.0082
	[0.0296]	[0.0286]	[0.0295]	[0.0335]	[0.0346]	[2.5152]	[0.0332]	[0.0334]
Shock, users	0.0383	0.0431		0.0815**	0.0660	6.0644		
	[0.0379]	[0.0377]		[0.0414]	[0.0454]	[3.7894]		
Shock, nonusers	–0.0548	–0.0307		–0.0844*	–0.0641	–2.6646		
	[0.0440]	[0.0400]		[0.0508]	[0.0484]	[3.0531]		
Shock, non–users | user Xs	–0.0922*	–0.0802		–0.1655***	–0.1720***	–7.7814*		
	[0.0534]	[0.0530]		[0.0636]	[0.0645]	[4.4042]		

Notes: The first three columns report the results from reweighting the data as per Fitzgerald, Gottschalk, and Moffitt (1998).

Table A20: IV results (cross-section and panel), a reproduction of Jack and Suri's Table 9

	Total consumption						Pr [Receive]	
	Cross-section	Panel	Panel	Panel without Msa	Panel	Panel without Msa	Panel without Msa	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
M-PESA user								
	–0.4705*	–0.5128***	–0.6328***	–0.6730***	–0.3155	–0.2561	–0.1455	0.0849
	[0.2685]	[0.1885]	[0.1851]	[0.2039]	[0.8549]	[1.5967]	[0.1741]	[1.1219]
Negative shock	–0.3344**	–0.3601**	–0.3462**	–0.4372**	–0.3762**	–0.4549**	–0.1739	–0.1395
	[0.1469]	[0.1668]	[0.1602]	[0.1741]	[0.1547]	[0.2095]	[0.1584]	[0.1718]
User × shock	0.5124*	0.6146**	0.5992**	0.7619**	0.6782**	0.8180**	0.3862	0.3080
	[0.2653]	[0.2908]	[0.2771]	[0.3084]	[0.2679]	[0.3655]	[0.2732]	[0.2903]
Controls	Yes		Yes	Yes	Yes	Yes	Yes	Yes
Rural × time FE					Yes	Yes		Yes
Location + Rural FE	Yes							
Observations	3,911	3,926	3,894	3,688	3,894	3,688	3,688	3,688
Shock effect	–0.0519**	–0.0217	–0.0159	–0.0204	–0.0024	–0.0074	0.0373	0.0290
	[0.0264]	[0.0365]	[0.0346]	[0.0359]	[0.0314]	[0.0333]	[0.0302]	[0.0313]
Shock, users	0.1781	0.2545*	0.2530**	0.3247**	0.3020**	0.3631**	0.2123*	0.1685
	[0.1237]	[0.1331]	[0.1255]	[0.1427]	[0.1205]	[0.1618]	[0.1212]	[0.1246]
Shock, nonusers	–0.3344**	–0.3601**	–0.3462**	–0.4372**	–0.3762**	–0.4549**	–0.1739	–0.1395
	[0.1469]	[0.1668]	[0.1602]	[0.1741]	[0.1547]	[0.2095]	[0.1584]	[0.1718]
Kleibergen–Paap LM test	39.796	46.681	48.243	45.244	3.732	1.734	45.244	1.734
LM test *p*-value	0.0000	0.0000	0.0000	0.0000	0.2919	0.6293	0.0000	0.6293
Kleibergen–Paap *F* statistic (critical value =11)	9.390	12.346	12.706	12.568	0.926	0.426	12.568	0.426

Table A21: Summary statistics (period 2) by adoption status (full sample), a reproduction of Jack and Suri's Online Appendix Table 1 (updated version)

	Early adopter		Late adopter		Non-adopter	
	Mean	SD	Mean	SD	Mean	SD
Own cell phone	0.940	0.237	0.885	0.320	0.368	0.483
Per capita consumption	87728	110733	57333	70384	38371	53414
Per capita food consumption	35627	27361	28948	24967	23558	22295
Total wealth	220859	1013048	107213	472330	58484	228156
HH size	4.278	2.225	4.737	2.398	4.252	2.384
Education of head (years)	8.673	5.341	7.683	4.667	5.611	4.366
Positive shock	0.075	0.263	0.076	0.266	0.050	0.218
Negative shock	0.604	0.489	0.526	0.500	0.578	0.494
Weather/Agricultural shock	0.134	0.341	0.114	0.319	0.146	0.354
Illness shock	0.443	0.497	0.361	0.481	0.415	0.493
Send remittances	0.660	0.474	0.505	0.500	0.167	0.373
Receive remittances	0.556	0.497	0.485	0.500	0.175	0.380
Financial access dummies						
Bank account	0.733	0.443	0.521	0.500	0.184	0.388
Mattress	0.679	0.467	0.744	0.437	0.857	0.351
Saving and credit cooperative	0.245	0.431	0.163	0.369	0.098	0.298
Merry go round/ROSCA	0.533	0.499	0.453	0.498	0.372	0.484
Household head occupation dummies						
Farmer	0.169	0.375	0.243	0.429	0.461	0.499
Public service	0.056	0.230	0.033	0.178	0.004	0.067
Professional occupation	0.236	0.425	0.223	0.416	0.102	0.303
Househelp	0.113	0.317	0.122	0.327	0.066	0.249
Run a business	0.177	0.382	0.144	0.351	0.166	0.373
Sales	0.112	0.315	0.099	0.299	0.052	0.221
In industry	0.024	0.152	0.013	0.115	0.019	0.137
Other occupation	0.038	0.192	0.050	0.219	0.040	0.196
Unemployed	0.071	0.258	0.072	0.259	0.082	0.275
Number of observations	1007		669		516	

Table A22: Risk sharing controlling for remittances, dependent variable is total consumption, a reproduction of Jack and Suri's Online Appendix Table 2

	(1)	(2)	(3)	(4)	(5)	(6)
	Overall shock		Illness shock		Illness shock	
	Original spec	Control for remittances	Original spec	Control for remittances	Original spec	Control for remittances
M-PESA user	0.0020	0.0153	0.0386	0.0561	0.0618	0.0674
	[0.0470]	[0.0477]	[0.0434]	[0.0446]	[0.0434]	[0.0443]
Negative shock	0.1544	0.1420	–0.0260	–0.0771	–0.0104	–0.0501
	[0.1627]	[0.1647]	[0.1589]	[0.1574]	[0.1515]	[0.1519]
User × shock	0.1380**	0.0972	0.1585**	0.0961	0.0630	0.0086
	[0.0632]	[0.0639]	[0.0728]	[0.0754]	[0.0731]	[0.0747]
Controls	Y	Y	Y	Y	Y	Y
+ Interaction	Y	Y	Y	Y	Y	Y
Time FE	Y	Y	Y	Y	Y	Y
Time × location FE	Y	Y			Y	Y
Observations	3,911	3,911	3,911	3,911	3,911	3,911
R-squared	0.323	0.329	0.150	0.161	0.323	0.330
Shock effect	–0.0041	–0.0018	0.0466	0.0451	0.0367	0.0335
	[0.0294]	[0.0293]	[0.0331]	[0.0330]	[0.0320]	[0.0318]
Shock, users	0.0415	0.0410	0.1104***	0.1069**	0.0781*	0.0708*
	[0.0375]	[0.0373]	[0.0423]	[0.0421]	[0.0406]	[0.0402]
Shock, nonusers	–0.0601	–0.0543	–0.0316	–0.0309	–0.0142	–0.0123
	[0.0442]	[0.0440]	[0.0503]	[0.0498]	[0.0477]	[0.0475]
Mean of user	0.5512	0.5512	0.5512	0.5512	0.5512	0.5512
Mean of shock	0.5344	0.5344	0.3231	0.3231	0.3231	0.3231

Table A23: Risk sharing and savings for Western Province (rounds 3 and 4), a reproduction of Jack and Suri's Online Appendix Table 3

	(1) Total consumption	(2) Prob [receive]	(3) Total received (square root)	(4) Total savings	(5) Total consumption (square root)	(6) Log total savings
M-PESA user	–0.4685*	–0.0784	–7.8547	–2295.8*	–7.4844	0.2630
	[0.2598]	[0.1561]	[9.3663]	[1338.8]	[7.4713]	[0.4752]
Negative shock	0.5659	0.2420	–7.5592	1615.7	19.643	0.1677
	[0.5123]	[0.3009]	[20.719]	[2932.5]	[22.281]	[0.8300]
User × shock	0.5624**	0.3325*	22.705**	1801.7	9.8975	–0.1320
	[0.2779]	[0.1828]	[10.514]	[1309.5]	[8.2443]	[0.5201]
Controls	Y	Y	Y	Y	Y	Y
+ Interactions	Y	Y	Y	Y	Y	Y
Time FE	Y	Y	Y	Y	Y	Y
Time × location FE	Y	Y	Y	Y	Y	Y
Observations	359	359	355	333	336	310
R-squared	0.399	0.271	0.241	0.232	0.333	0.415
Mean of user	0.8094	0.8094	0.8070	0.8152	0.8161	0.8359
Mean of shock	0.5900	0.5900	0.5939	0.5807	0.5820	0.5901

Table A24: Basic three-period results, a reproduction of Jack and Suri's Web Appendix Table 1

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Overall shock		Rainfall shock	Illness shock		Overall shock	
		Total consumption		Non-med consumption	Prob [receive]	Number received	Total received (root)
M-PESA user	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Shock	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
User × shock	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Controls	Y	Y	Y	Y	Y		Y
+ Interactions	Y	Y	Y	Y	Y		Y
Time FE		Y			Y		Y
Time × location FE		Y					Y
Observations	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
R-squared	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Negative shock effect	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Shock, users	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Shock, nonusers	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Table A25: Reduced forms using agent rollout: overall shock, a reproduction of Jack and Suri's Web Appendix Table 2

	(1) Agents within 1 km	(2) Agents within 1 km	(3) Agents within 2 km	(4) Agents within 5 km	(5) Agents within 20 km	(6) Distance to closest agent	(7) Distance to closest agent
Shock	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Agent variable	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Agent × shock	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Controls	Y	Y	Y	Y	Y	Y	Y
+ Interactions	Y	Y	Y	Y	Y	Y	Y
Time FE		Y	Y	Y	Y		Y
Time × location FE		Y	Y	Y	Y		Y
Observations	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
R-squared	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mean of agents	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Appendix B: Proposed but undone robustness checks

Since my analysis plan was written with minimum interaction with the data, I was unable to confirm with certainty whether I am eventually able to obtain the data. These robustness checks are those I initially planned to implement but were unable to complete, however, due mainly to the unavailability of data.

Placebo test

My next approach considered running a placebo test. The idea behind the placebo test is that different outcomes between M-PESA users and nonusers could be due to unobserved factors that were correlated with the adoption of M-PESA. My planned approach was to analyse the period immediately preceding the introduction of M-PESA and compare outcomes between 'future M-PESA users' and 'future M-PESA nonusers'. If I observe differences between these two groups in the period before the advent of M-PESA that suggests that subsequent differences might be due to factors other than M-PESA. However, if no differences are observed in the pre-M-PESA period that supports the idea that subsequent differences are, in fact, due to the introduction of M-PESA.

My planned approach was to use data from a four-period panel household agricultural survey collected by the Tegemeo Institute over the years 1997–2007 (Tegemeo n.d.),[25] which preceded the introduction of M-PESA. I would create a dummy variable for the households that would later use M-PESA after its introduction in 2007 and then estimate the following specification:

$$C_{ijt} = \alpha_i + \gamma Shock_{ijt} + \mu MPUser_{ijt} + \beta MPUser_{ijt} \times Shock_{ijt} + \theta^S X_{ijt} \times Shock_{ijt} + \theta^M X_{ijt} + \eta_{jt} + \pi_{rt} + \varepsilon_{ijt}\,, \qquad (A-1)$$

where $MPUser_{ijt}$ is a dummy variable which takes the value of one if a household later used M-PESA and zero otherwise. If the estimated coefficient for $MPUser_{ijt}$ were insignificant this would support the hypothesis that the observed M-PESA effect was real and not due to unobservable characteristics of households or the location of agents.

Unfortunately, I was unable to run the test due to the unavailability of data. I also tried to use FinAccess 2006, 2009 and 2013 surveys. However, the data on shocks were not collected in those surveys.[26]

Heterogeneous slopes

In order to address the potential bias resulting from unobserved, time-invariant household heterogeneity, the fixed effects model has been estimated throughout the original paper. However, as a further robustness check, I planned to allow for heterogeneous individual specific slope on the shock variable to examine to what extent

[25] The household survey data for 2006 has been recently publicly available (http://fsdkenya.org/dataset/finaccess-household-2006/). Thus, I can see which of these two data sets may help me to conduct this test.

[26] FinAccess household surveys for several years are publicly available via: https://dataverse.harvard.edu/dataverse.xhtml?alias=fsdkenya

the results on β, the coefficient on the interaction of interest, are robust to this change. My planned approach was to add an interaction between household fixed effects, α_i, and shock variable, $Shock_{ijt}$. The following specification control for such possibility:

$$C_{ijt} = \alpha_i + \gamma Shock_{ijt} + \mu User_{ijt} + \beta User_{ijt} \times Shock_{ijt} + \theta^S X_{ijt} \times Shock_{ijt} + \theta^M X_{ijt} + \omega \alpha_i Shock_{ijt} + \varepsilon_{ijt} \quad \text{(A – 2)}$$

In other words, I would control for any effects unobservable time-invariant household specific may have had on the ability of households to smooth income shocks. However, the major concern to regress the above mention specification is that the associated estimates might be biased due to the short time series available in the original study (for example, three rounds of panel data). To avoid such a potential bias, my planned approach was to use five rounds of household panel survey data over the years 2008 and 2014 conducted by the original authors after this study for another research paper (Jack and Suri 2016). Unfortunately, given that the T is small and there is no access to the follow-up surveys, I was unable to conduct this test.

References

Aker, JC and Mbiti, IM, 2010. Mobile phones and economic development in Africa. *The Journal of Economic Perspectives*, 24(3), pp.207–232.

Alderman, H and Paxson, CH, 1992. Do the poor insure? A synthesis of the literature on risk and consumption in developing countries. *World Bank Publications*, 164.

Alinaghi, N and Reed, WR, 2017. *Risk sharing and transaction costs: evidence from Kenya's mobile money revolution, a replication study of Jack and Suri (2014)*. Available at: < http://3ieimpact.org/evidence-hub/replication-studies-status/risk-sharing-and-transactions-costs-evidence-kenyas-mobile >

Brown, AN, Cameron, DB and Wood, BD, 2014. *Quality evidence for policymaking: I'll believe it when I see the replication*. 3ie Replication Paper 1. Washington, DC: International Initiative for Impact Evaluation (3ie).

Cox, N, 2004. Extremes: Stata module to list extreme values of a variable. *Statistical Software Components*. Boston College Department of Economics.

Cragg, JG and Donald, SG, 1993. Testing identifiability and specification in instrumental variable models. *Econometric Theory*, 9(2), pp.222–240.

De Weerdt, J and Dercon, S, 2006. Risk-sharing networks and insurance against illness. *Journal of Development Economics*, 81(2), pp.337–356.

Fafchamps, M and Lund, S, 2003. Risk-sharing networks in rural Philippines. *Journal of Development Economics*, 71(2), pp.261–287.

Fitzgerald, J, Gottschalk, P, and Moffitt, R. A., 1998. An analysis of sample attrition in panel data: The Michigan Panel Study of Income Dynamics. *Journal of Human Resources*, 33(2), pp.251–299.

GSMA, 2017. *2017 Mobile Industry Impact Report: Sustainable Development Goals*. Available at: <https://www.gsmaintelligence.com/research/?file=f62a94c22f3f7e90fc2165a8dd5b44bc&download=&utm_source=website&utm_medium=button&utm_campaign=Track_Downloads&utm_term=SDGs&utm_content=PDF> [Accessed 18 December 2018].

Jack, W and Suri, T, 2011. *Mobile money: the economics of M-PESA (No. w16721)*. National Bureau of Economic Research.

Jack, W and Suri, T, 2014. Risk sharing and transactions costs: evidence from Kenya's mobile money revolution. *The American Economic Review*, 104(1), pp.183–223.

Kleibergen, F and Paap, R, 2006. Generalized reduced rank tests using the singular value decomposition. *Journal of Econometrics*, 133(1), pp.97–126.

Mbiti, I and Weil, DN, 2011. *Mobile banking: the impact of M-PESA in Kenya (No. w17129)*. National Bureau of Economic Research.

Munyegera, GK and Matsumoto, T, 2016. Mobile money, remittances, and household welfare: panel evidence from rural Uganda. *World Development*, 79, pp.127–137.

Osborne, JW and Overbay, A, 2004. The power of outliers (and why researchers should always check for them). *Practical Assessment, Research and Evaluation*, 9(6), pp.1–12.

Stock, JH, Wright, JH and Yogo, M, 2002. A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business and Economic Statistics,* 20(4), pp.518–529.

Stock, JH and Yogo, M, 2005. Testing for weak instruments in linear IV regression. *In: Andrew DWK identification and inference for Econometric Models*. New York: Cambridge University Press, pp. 80-108.

Suri, T and Jack, W, 2016. The long-run poverty and gender impacts of mobile money. *Science*, 354(6317), pp.1,288–1,292.

Tegemeo, n.d. *Open Data.* Tegemeo Institute of Agricultural Policy and Development, Egerton University. Available at: <http://www.tegemeo.org/index.php/resources/data/230-request-for-data.html> [Accessed 18 December 2018].

UNCDF, 2018. Financial Inclusion and the SDGs. Available at: <http://www.uncdf.org/financial-inclusion-and-the-sdgs> [Accessed 18 December 2018].

Williams, R, 2016. Outliers. University of Notre Dame. Available at: <https://www3.nd.edu/~rwilliam/stats2/l24.pdf> [Accessed 18 December 2018].

Other publications in the 3ie Replication Paper Series

The following papers are available from http://3ieimpact.org/evidence-hub/publications/replication-papers

Cash and change: a replication study of a cash transfer experiment in Malawi, 3ie Replication Paper 21. Reimão, ME, 2018.

Impact of unconditional cash transfers: a replication study of the short-term effects in Kenya. 3ie Replication Paper 20. Wang, H, Qui, Fang, Q and Luo, J, 2018.

Mobile money and its impact on improving living conditions in Niger: a replication study. 3ie Replication Paper 19. Meneses, JP, Ventura, E, Elorreaga, O, Huaroto, C, Aguilar, G, and Beteta, E, 2018.

Savings revisited: a replication study of a savings intervention in Malawi, 3ie 3ie Replication Paper 18. Stage, J and Thangavelu, T 2018.

Thou shalt be given…but how? A replication study of a randomized experiment on food assistance in northern Ecuador, 3ie Replication Paper 17. Lhachimi, SK and Seuring, T, 2018.

Preventing HIV and HSV-2 through improving knowledge and attitudes: a replication study of a multicomponent intervention in Zimbabwe. 3ie Replication Paper 16. Hein, NA, Bagenda, DS and Yu, F, 2018.

PEPFAR and adult mortality: a replication study of HIV development assistance effects in Sub-Saharan African countries. 3ie Replication Paper 15. Hein, NA, Bagenda, DS and Luo, J, 2018.

When to start ART? A replication study of timing of antiretroviral therapy for HIV-1-associated Tuberculosis. 3ie Replication Paper 14. Djimeu, EW, 2018.

STRETCHing HIV treatment: a replication study of task shifting in South Africa. 3ie Replication Paper 13. Chen, B and Alam, M, 2017.

Cash transfers and HIV/HSV-2 prevalence: a replication of a cluster randomized trial in Malawi. 3ie Replication Paper 12. Smith, LM, Hein, NA and Bagenda, DS, 2017.

Power to the people?: a replication study of a community-based monitoring programme in Uganda, 3ie Replication Paper 11. Donato, K and Garcia Mosqueira, A, 2016.

Fighting corruption does improve schooling: a replication study of a newspaper campaign in Uganda, 3ie Replication Paper 10. Kuecken, M, and Valfort, MA, 2016.

The effects of land titling on the urban poor: a replication of property rights, 3ie Replication Paper 9. Cameron, Drew B, Whitney, Edward M and Winters, Paul C, 2015.

Male circumcision and HIV acquisition reinvestigating the evidence from young men in Kisumu, Kenya, 3ie Replication Paper 8. Djimeu, EW, Korte, JE and Calvo, FA, 2015.

Walking on solid ground: a replication study on Piso Firme's impact, 3ie Replication Paper 7. Basurto, MP, Burga, R, Toro, JLF and Huaroto, C, 2015.

The impact of India's JSY conditional cash transfer programme: A replication study, 3ie Replication Paper 6. Carvalho, N and Rokicki, S, 2015.

Recalling extra data: A replication study of finding missing markets, 3ie Replication Paper 5. Wood, BDK and Dong, M, 2015.

The long and short of returns to public investments in fifteen Ethiopian villages, 3ie Replication Paper 4. Bowser, WH, 2015.

Reanalysis of health and educational impacts of a school-based deworming program in western Kenya Part 2: Alternative analyses, 3ie Replication Paper 3, part 2. Aiken, AM, Davey, C, Hayes, RJ and Hargreaves, JR, 2014.

Reanalysis of health and educational impacts of a school-based deworming program in western Kenya Part 1: A pure replication, 3ie Replication Paper 3, part 1. Aiken, AM, Davey, C, Hargreaves, JR and Hayes, RJ, 2014.

TV, female empowerment and demographic change in rural India, 3ie Replication Paper 2. Iversen, V and Palmer-Jones, R, 2014.

Quality evidence for policymaking: I'll believe it when I see the replication, 3ie Replication Paper 1. Brown, AN, Cameron, DB and Wood, BDK, 2014.

Replication Paper Series

International Initiative for Impact Evaluation
1029 Vermont Avenue, NW
Suite 1000
Washington, DC 20005
USA

replication@3ieimpact.org
Tel: +1 202 629 3939

www.3ieimpact.org