A Replication study: Payment Mechanisms and Antipoverty Programs: Evidence from a Mobile Money Cash Transfer Experiment in Niger. Aker et al. (2016)

Submitted as a replication proposal to 3ie Replication Window 4

I. Introduction

Mobile money is a revolutionary service in which the mobile phone is used to access financial services[1]. This new transference mechanism has lower fees and easier use than traditional payment systems, resulting a great opportunity for vulnerable groups. According to Mbiti et al. (2011), M-Pesa, a mobile money service, reduced prices of competing money transfer services such as Western Union and decreased informality for savings, facts that improved population welfare. In addition, farmers in Niger recognized time savings for each payment are equivalent to an amount that would feed a family of five for a week. Digital payments increase control, since senders of remittances can have a greater influence on how recipients use the money, including for savings, Grossman et al. (2014).

As a consequence of these promising results, the volume of mobile payments on M-Pesa in Kenya was about US$24 billion[2]. Besides, in 2013, 66% of Kenyans said their mobile money accounts are very important in their finances[3]. This to get a sense of the magnitude and utility of digital payment flows.

Niger is one of the poorest countries in the world. Between 2000 and 2015 the Gross National Income per capita in Niger raised since $170 to $390[4], even remaining as the fifth lowest income in the world. Additionally, there is much work remains to be done in improving quality of life for people, the percentage of the population living below the national poverty line was 49% in 2011[5]. Besides, the country had important problems on education; Adult literacy rate had an important decreased from 28.7 in 2005 and 15.5 in 2012[6].

On the other hand, the mobile cellular subscriptions (per 100 people) had a significant increased between 2.5 in 2005 to 44.4 in 2014. Furthermore, population covered by a mobile-cellular network upgrade since 15% to 75%, during the same period[7]. This fact represents a great opportunity for this country in order to develop mobile money and all its potential benefits.

Paper description

The study of Aker et al. (2016) used a randomized experiment to evaluate the effects of using mobile money as cash transfer program in Niger. In this order, vulnerable households in 96 villages, represented by woman, were selected to receive an unconditional cash transfer through three delivery channels. These three different interventions were *Cash, Mobile* and *Zap*, each of them with an equal amount of randomized-selected villages. All beneficiaries received a cash transfer of CFA 22,000 (approximately $45) per month during five periods. *Cash* program recipients obtained an individual envelop with money, *Mobile* intervention included the same features as *Cash* plus a m-transfer-enabled mobile phone. The *Zap* recipients used the m-transfer system in order to get the cash as well as m-transfer-enabled mobile phone.

The original paper results provide evidence that *Zap* intervention had a superior impact than *Cash* and *Mobile*. The main findings of Aker et al. (2016) are the increase in consumption of beans and fats in 30% and improvement about diversity of food and nonfood items, 0.78 and 0.85 more than *Cash* and *Mobile* respectively (significant at the 1% level), these are particularly important given the high prevalence of protein-energy malnutrition in Niger (INS 2013).

[1] Mobile Money definitions – GSMA 2010
[2] Bank of England 2014; https://blockchain.info; company reports.
[3] FII Kenya 2013.
[4] World Bank Data Base (2016), GNI per capita, Atlas method (current US$).
[5] World Bank, Global Poverty Working Group.
[6] UNESCO Institute for Statistics (2016)
[7] The Little Data Book on Information and Communication Technology 2015. World Bank (2015)

The following suggested channels supported these results: Reduced costs of obtaining the transfer and reforming intra-household dynamics. On the other hand, the intervention would decrease the time of the cash transfer.

Policy relevance

According to World Bank, one of the causes of poverty is vulnerability to negative shocks. The deeper cause is the inability to reduce or mitigate risk or cope with shocks. In these sense, effective tools for saving, sending, and borrowing money and mitigating financial risks can help people to go out of poverty.

For example, access to credit prevents the unnecessary depletion of capital by poor producers who do not have sufficient reserves to face an unexpected negative shock, Rosenzweig & Wolpin (1993). Additionally, the systematic review of Pande et al. (2012) shows compelling evidence that the poor people's access to formal banking services can raise their incomes. Access to formal banking services produces information ex ante about possible investments and allocation of capital. Besides, banking technology like mobile phones can facilitate savings, remittances, transfers and payments among the poor and increase incomes by allowing households to smooth consumption and accumulate savings.

For that reason, Bill and Melinda Gates foundation believe that connecting poor people with digitally-based financial tools and services can improve their quality of life. In this order, the study of Aker et al. (2016) represents a milestone in public policy research related to mobile-money transfer programs.

According to World Food Program, the malnutrition rates in Niger are elevated; ten percent of children under five suffer from acute malnutrition and 44 percent of children suffer from chronic malnutrition. Also, 73% of children aged 6-59 months are anemic (INS 2013). Aker et al (2016) provides evidence about the potential impact of using m-transfer system to increase diet diversity. This could be a channel to solve these problems, particularly nowadays.

Currently, the possibilities to use mobile money had increased in Niger. In 2015, the MFS Africa Hub connects mobile wallet customers across networks and countries in Sub-Saharan Africa. This means that Niger is able to send and receive e-money to or from Benin and Cote d'Ivore, the principal sources of remittances after Nigeria. It is expected that this increased the welfare of people because a money transfer through mobile money are instantaneous, secure, traceable, and dramatically cheaper in contrast with others formal and informal channels. Besides, according to GSMA, the mobile subscriptions and connections in Sub-Saharan Africa are predicted to rise substantially between 2013 and 2020, 62% in number of subscribers and 68% in number of connections.

II. Replication Plan

2.1. Pure Replication

This section presents a reproduction of results obtained from Aker et al. (2016). According to Hamermesh (2007) and Brown et al (2014), a pure replication is the reproduction of the original study using the original data in order to validate the original results. This includes reconstructing all variables, writing and running the programs independently. For this porpoise, we keep the same assumptions and methodology used by Aker et al (2015). In this sense, we expect to access the four primary sources of information used by the authors. The first data set is composed of three household surveys of program recipients (baseline in May 2010, follow-ups in December 2010 and May 2011. The second is a village-level survey, collected from focus group. The third is weekly price information for six products in 45 markets between May 2010 and January 2011. The last is anthropometric data among children under 5 years old, collected in May 2011.

On the other hand, we consider performing a double-checking of all the code, especially for the procedures that are not part of the statistical package used by original authors. In addition, we will include an independent outlier analysis and the decisions taken in relation to its results.

Finally, this section consider that we will be able to access all documentation in order to get a wider approach to the sample design, the process of producing the database and the data dictionaries with the purpose of understanding each variable and additional routines to cleaning the data.

2.2. Measure and Estimation Analysis (MEA)

Missing value imputation

We propose to apply Multiple Imputation (MI) methodology to address the missing data problem in order to test the robustness of the results on the original paper. Multiple Imputation is widely accepted as the “state of the art” method for imputation (Enders, 2010 and Schafer et al. 2002). Currently, it has received considerable attention in epidemiology, psychology and political science and sociology.[8]

Original authors do not report using any particular strategy for dealing with missing data, therefore, we can assume they use the method of “listwise deletion” or “pairwise deletion”, which involves only using for the analysis the observations where they have complete information for all variables[9]. According to the literature (Enders, 2010 and Schafer et al. 2010), this method might provide less precise estimates and even biased estimations if data is no “Missing Completely at Random”(MCAR).

As mentioned by Enders (2010), the use of multiple imputation on applied social research is increasing, but it is still far from being commonly used, mainly due to two factors. One is that statistical packages have as default other imputation strategies, the other is that making a correct imputation implies more sophisticated use of statistics, which can make researchers redundant of using it as can be seen as not transparent, in comparison with more simple methods that use the default option (which is often, pairwise deletion).

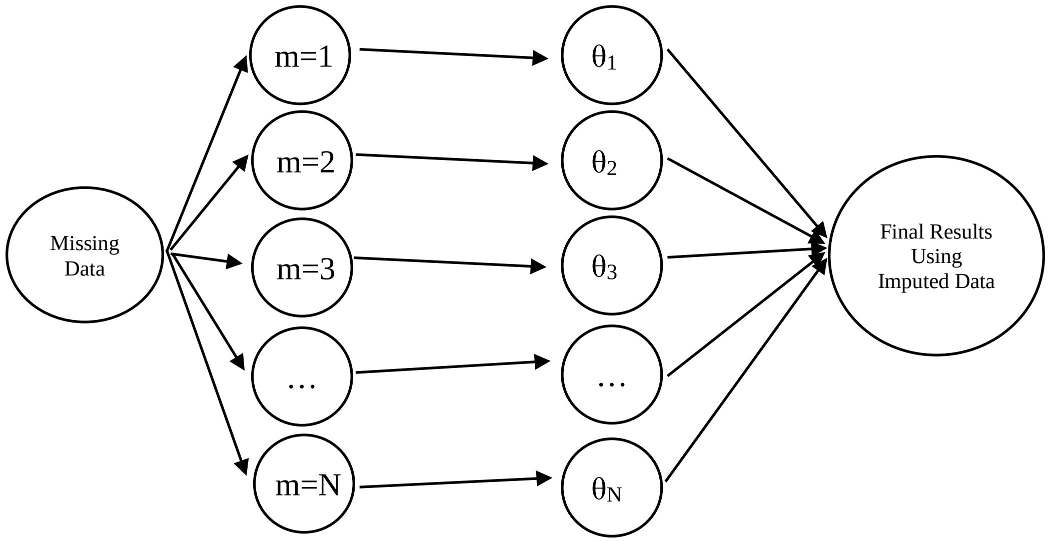
The general idea of MI can be summarized in Figure 1, where we see that it implies three phases: i) Imputation, ii) Analysis and iii) Pooling. In the first phase, we need to select an imputation algorithm and choose the number of imputations. We will select the method of Predictive Mean Matching (PMM), and we will select the covariates that could not be affected by the treatment and that have high predictive power.

Figure 1. Graphical description of Multiple Imputation analysis

Phases i) Imputation phase ii) Analysis phase iii) Pooling phase

[8] As an example, the recent working paper of Lall (2016) and Breitwieser and Wick (2016) test the robustness of previous results on political economy and find that using MI substantially changes the interpretation of the revised studies.

[9] Is not possible to determine which one based on the paper, but this are the only two possible options that can be inferred.



Based on Enders (2010).

The second is the analysis phase where we use the results to obtain our estimation of interest (for example, regression analysis) for each imputation performed. It is important to remark, that the reason for the differences between imputations is the addition of an error term that allows the variability in the estimation. This is a common feature in Bayesian statistics, and is the crucial difference with other imputation methods which use only a single imputation.

Finally, the pooling phase uses the "Rubin rules" (Rubin, 1987) to combine all the information from the imputations to a single estimate. In general, it can be said the MI uses more efficiently all the information provided by the multiple imputations in order to provide more accurate estimates.

Attrition and missing data on the outcome

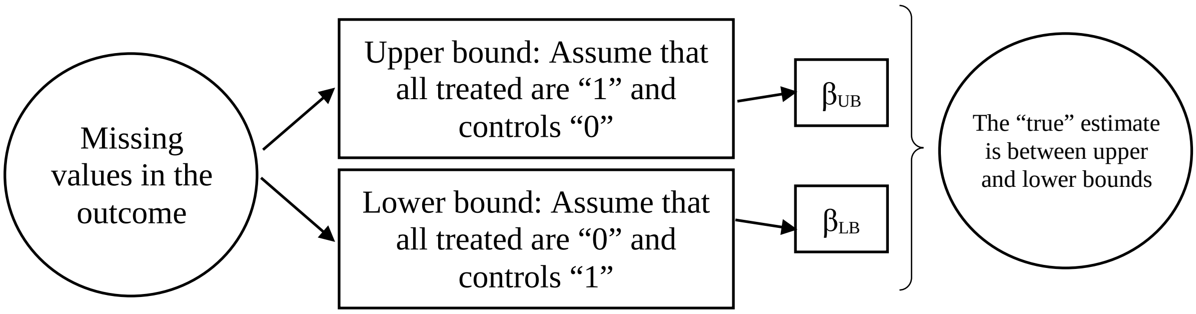
Authors report having an overall attrition of around 7%, which is an opportunity for further analysis[10]. Therefore, we propose to employ bounded support strategies (Horowitz and Manksi 2000, Lee 2009) to test the robustness of the original results obtained. Bounded analysis is a widely accepted method for dealing with attrition bias.

The main objective of this bounding analysis is to have additional robustness checks to the results and to check if the main conclusions of the paper hold with this method. It is important to mention that we consider that it is highly probable that only the upper bound will keep statistically significant (where we assume the best possible scenario for the imputation), as is usual in the literature and since most of the results that are statistically significant at the 5 or 10 percent of confidence. Nevertheless, we still find the robustness check informative.

The main idea of Horowitz and Manski (2000) extreme bound methodology is to estimate two bounds assuming extreme scenarios. For simplicity, we assume a binary outcome variable and that the impact of the treatment is to increase this outcome (positive relation). It can be summarized on Figure 2.

Figure 2. Graphical description of Extreme Bound Analysis

[10] For example, Gertler et al. (2011) mention that "Best-practice impact evaluations aim to keep nonresponse and attrition below 5 percent".



Based on Horowitz and Manski (2000).

The upper bound refers an “over optimistic” situation and assuming that all the missing values of the outcome are “1” for all treated individuals, and “0” for all control individuals. In the lower bound, it is done the opposite, supposing “over pessimistic” condition in which all the missing values in the outcome for the treated are “0” and “1” for the controls.

Nevertheless, as Lee (2009) points out, this method tends to offer large bounds; consequently, he proposed to use “Sharp bounds”, which are essentially a fusion of Horowitz and Manski (2000) method with Heckman self-selection methodology. In other words, Lee (2009) proposed to “group” cases on observable variables and then to impute the extreme values described.

Heterogeneity Results: Age and education

The author presents evidence on the channels through which the observed impact occurred. One of them is “increased use of mobile phones”. In theory, this channel should have a similar effect for Zap and Mobile groups. However, results show that cellphone usage was higher among Zap households compare with the Mobile households. This could be because Zap program recipients felt a greater sense of “Ownership”. Moreover, we consider that the age of programs recipients is also an important factor to explain the use of mobile phone and the managed of new technology. Furthermore, Mbiti et al. (2011) observe M-Pesa users are more likely to be younger, wealthier and better educated. These suggest that the impact of m-transfer could be bigger for younger or better educated. According to World Bank, there is a significant gap of 7.6 pp in the literacy rates on 2015 favoring young adults from 15 to 24 years (26.6%)[11]in contrast to people older than 15 year old (19.0%) in general.

In these sense, we are going to add interaction terms to the regression model in order to expand understanding of the relationships among the treatment, age and level of education. For that, we propose to use a STATA® command called *“hte”* to analyze the treatment-effect heterogeneity. The approach of *“hte”* is to assume, at least provisionally, conditional unconfoundedness given a set of covariates and use propensity score stratification to estimate treatment effects at various points over the range of the propensity score, Becker, S. O. and Ichino, A. (2002).

2.3 Theory of Change Analysis (TCA)

Nutrition Evaluation

Every day there is more evidence about the risk factor that implies malnutrition for many etiologies and how it increases mortality, besides the relevance of feeding to assess a real impact in childhood growth as part of public policy intervention. Moreover, a deeper analysis is especially appropriate for this study because of the diet diversity improvements found for Zap group, without statistical differences when nutrition status was evaluated (Z-score means).

[11] Source: UNESCO Institute for Statistics.

In order to amplify the nutrition evaluation performed by the original authors, we will add complementary analysis to weight-for-height z-score in children younger than 5 years, a widely accepted evaluation established by WHO (2006).

First, we will start defining wasting diagnosis, which indicates acute malnutrition.

Wasting:

- Moderate wasting: weight-for-height Z-Score falling < -2 to -3 SD
- Severe wasting : weight-for height Z-Score falling bellow -3 SD

The additional use of anthropometric data 0 to 5 years-old period available will be useful in order to get complementary scores also described by WHO, in specific stunting, which correlates with chronic malnutrition.

Stunting:

- Moderate stunting – height-for-age Z-score falling <-2 to -3 SD
- Severe stunting – height-for-age Z-score falling below -3 SD

After this initial analysis, we will have a proportion of people in each group with presence or absence of underweight, and its specific classification: chronic moderate or severe, acute moderate or severe and not malnourished.

With these proportions, it will be analyzed potential associations of higher or less frequencies of nutritional diseases in a specific intervention group: Zap, Mobile and Cash. Consequently, we will perform a bivariate analysis Zap-Cash, Cash-Mobile and Zap-Mobile, using cases and controls of our nutritional diagnosis of interest in order to get Chi-square outcome with statistical significance. In other words, this analysis will help us to establish a potential association with statistical significance of moderate stunting, severe stunting, moderate wasting and severe wasting with Zap, Cash or Mobile groups.

Furthermore, it would be relevant to approximate a statistical analysis for identify the age benefitting the most by this intervention. For this purpose, we include a heterogeneity analysis by age to evaluate differences on wasting and stunting diagnosis in children less than five years.

III. Conclusion

This replication study aims to validate the findings in Aker et al. (2016) and explore the conclusions about the impact of m-transfer mechanism in Niger. In this order, we first conduct a pure replication of the original results and validating the model assumptions using by the original authors. We also propose to employ Multiple Imputation (MI) methodology to address the missing data problem in order to test the robustness of the results. Additionally, we prove if the impact on diet diversity is higher in younger, more educated or household with children under 5 years old. Finally, we amplify the nutrition evaluation performed by the original authors. Through these activities, we will generate complementary analysis based on the original research.

Conflict of Interest Disclosure: The authors do not present any conflict of interest in relation to the present project. We have not contacted or interacted previously with the original data/code from the original authors. In addition, the PIs or Co-PIs ha have not directly worked or contacted the researchers from the original analysis. We have not receive payment or services from a third party (government, commercial, private foundation, etc.) for any aspect of the submitted work.

References

-Aker, J, Boumnijel, R, McClelland, A, and Tierney, N, (2016). Payment mechanism and anti-poverty programs: evidence from a mobile money cash transfer experiment in Niger. Economic Development and Cultural Change. Forthcoming.

-Aker, Jenny C., Christopher Ksoll, and Travis J. Lybbert. (2010). "ABC, 123: The Impact of a Mobile Phone Literacy Program on Educational Outcomes." CGD WP. 223. Washington. D.C.

-Aker, Jenny C. and Isaac M. Mbiti. (2010). "Mobile phones and economic development in Africa." Journal of Economic Perspectives, 24(3): 207–232.

-Bank of England. (2014). "Innovations in Payment Technologies and the Emergence of Digital Currencies." Quarterly Bulletin 2014 Q3.

-Beal, G., Rogers, E., and Bohlen, J. (1957). Validity of the concept of stages in the adoption process. *Rural Sociology*, 22(2):166–168.

- Becker, S. O. and Ichino, A. (2002). Estimation of average treatment effects based on propensity scores. The Stata Journal 2:358–377.

-Bharat, T. (2002). Technological Change and the Dispersion of Wages». FRBSF Economic Letter, 2002-23. 2002.

-Bohlen, J., and Beal, G. (1957). The Diffusion Process. Special Report No. 18, Agriculture Extension Service, Iowa State College, 1: 56–77.

-Breitwieser, A. and Wick, K. (2016). What We Miss By Missing Data: Aid Effectiveness Revisited. *World Development*, 78(1): 554-71.

-Brown, A. N., Cameron, D. and Wood, Benjamin DK. (2014). Quality evidence for policymaking: I'll believe it when I see the replication. 3ie Replication Paper 1

-Brynjolfsson, E. and Yang, S. (1996). Information Technology and Productivity: A Review of the Literature. *Advances in Computers*, 43, 179:214.

-Buschi, F., Fanzo, J. and Frinzo, E. The Rule of Food and Nutrition System Approaches in Tackling Hidden Hunger (2011) Int J Environ Res Public Health, 8(2): 358–373.

-Card, D. and Di Nardo, D. (2002). Skill biased Technological change and rising Wage Inequality: Some problems and puzzles». NBER Working Paper 8769.

- De Sardan, Olivier, J. P, Hamani, O., Isaley, N., Y, Adamou, H. and Oumarou, I. (2014) Cash transfer in Niger: the manna, the norms and the suspicions.

-Enders, C. (2010). Applied Missing Data Analysis. New York: The Guilford Press. Vienna Economics Papers 1302, University of Vienna, Department of Economics.

-FAO. (1985). Energy and Proteins Requirements. FAO Corporate Document Repository. Geneva

-Fanzo, J. (2012). The Nutrition Challenge in Sub-Saharan Africa. Regional Bureau for Africa Working Paper. United Nations Development Programme. 2012-012.

-Gertler, P., Martinez, S., Premand, P., Rawlings, L. and Vermeersch, C. (2011). Impact Evaluation in practice. Washington D.C.: The World Bank.

-Grossman, Jeremiah, and Michael Tarazi. (2014). Serving Smallholder Farmers: Recent Developments in Digital Finance. Washington, DC: CGAP.

-Hamermesh, DS. (2007). Viewpoint: replication in economics. The Canadian Journal of Economics, 40(3), pp.715–33

-Horowitz, J. and Manski, C. (2000). Nonparametric analysis of randomized experiments with missing covariate and outcome data. *Journal of the American Statistical Association*, 95: 77–84.

-INS (Institut National de la Statistique). (2013). "Enquête Démographique et de Santé et à Indicateurs Multiples du Niger, 2012." INS and ICF International, Calverton, MD.

-Lall, R. (2016). How Multiple Imputation makes a Difference. *Political Analysis.* Forthcomming.

-Lee, D. (2009). Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects. *Review of Economic Studies,* 76: 1071-1102.

-Legris, P., Ingham, J. and Collerette, P. (2003). Why do people use information technology? A critical review of the technology acceptance model. *Information & Management* 40(3), 191:204.

-Mbiti, Isaac M. and David Weil. (2011). "Mobile Banking: The Impact of M-Pesa in Kenya." NBER Working Paper No. w17129.

- Pande, R., Cole, S., Sivasankaran, A., Bastian, G., & Durlacher, K. (2012). Does poor people's access to formal banking services raise their incomes?. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

-Rogers, E. (1962). Diffusion of innovations (1st ed.). New York: Free Press of Glencoe.

-Rosenzweig, M. and Wolpin, K. (1993). “Credit Market Constraints, Consumption Smoothing and the Accumulation of Durable Production Assets in Low Income Countries: Investment in Bullocks India” Journal of Political Economy, 101, 2, pp, 223-244.
-Schafer, J. and Graham, J. (2002). Missing data: our view of the state of the art. *Psychological Methods,* 7(2): 147-77.
- UNESCO (2016). Institute for Statistics: UNESCO Institute for Statistics.
-Von Hippel, T. (2007). Regression with missing Ys: An improved strategy for analyzing multiply imputed data. Sociological Methodology, 37(1): 83-117.
-Vorster, J.; De Waal, L. (2008). “Beneficiaries of the Child Support Grant: Findings from a national survey”, in The Social Work Practitioner-Researcher, N° 20, pp.233-248.
-World Bank (2015). The Little Data Book on Information and Communication Technology 2015.
-WHO (2006). WHO Child Growth Standards: Length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: Methods and development. Geneva.
-WHO (2007). Protein and amino acid requirements in human nutrition: report of a joint FAO/WHO/UNU expert consultation. WHO technical report series; no. 935.