A Replication Plan for:

"Building State Capacity: Evidence from Biometric Smartcards in India"

Akinwande A. Atanda and W. Robert Reed

This replication plan is submitted for 3ie's Replication Window 4 on "Financial Services for the Poor"

Principal Investigator:

Akinwande A. Atanda

Department of Economics and Finance
University of Canterbury
Private Bag 4800
Christchurch
New Zealand
Phone: +64-22-4370774
Email: akinwande.atanda@pg.canterbury.ac.nz

Co-Investigator:

W. Robert Reed

Department of Economics and Finance
University of Canterbury
Private Bag 4800
Christchurch
New Zealand
Email: bob.reed@canterbury.ac.nz

Replicated Study:

Muralidharan, Karthik, Paul Niehaus and Sandip Sukhtankar. 2016. "Building State Capacity: Evidence from Biometric Smartcards in India." *American Economic Review*, 106(10): 2895-2929.

I. Introduction

As the world's second most populous country, India has experienced impressive economic gains over the past several decades. The percentage of Indians[1] living on less than \$1.90 (Rs. 126.5) and \$3.10 (Rs. 206.3) a day decreased by 80% and 27%, respectively, between 2004 and 2011 (World Bank, 2016). However, this growth has not been allocated evenly. The majority of the population that spends less than \$1.90 a day are located in rural areas, where the poverty rate is growing. This is reflected in the fact that only 32.6% of rural dwellers in India are account holders with access to formal financial institutions (World Bank, 2014). These statistics indicate that the poor are underserved by formal financial services such as credit, savings, insurance, and payments and remittance institutions. This financial exclusion problem has motivated interest in designing appropriate financial services that align with the needs of the poor in low-and middle-income countries such as India.

It is in this context that Karthik Muralidharan, Paul Niehaus and Sandip Sukhtankar (henceforth, MNS) published their study in 2016: "Building State Capacity: Evidence from Biometric Smartcards in India", *American Economic Review*, Volume 106, Number 10. MNS has been cited 44 times[2] (Google Scholar, 2017) and is one of the studies selected for 3ie's Replication Window 4. It reports the results of a large-scale experiment that randomized the rollout of a biometrically-authenticated payments infrastructure ("Smartcards") across 157 sub-districts in India. MNS investigated the impact of Smartcards on beneficiaries of employment (NREGS[3]) and pension (SSP[4]) programs in the Indian state of Andhra Pradesh (AP). The attraction of Smartcards is that they have the potential to enhance government's technical capacity to enable prompt payment transfers (Pritchett, 2010) and reduce the theft of money meant for the poor by government officials (i.e. leakages of funds) (Niehaus and Sukhtankar, 2013, Muralidharan, Niehaus, and Sukhtankar, 2014).

Replication of MNS has the potential to provide enhanced credibility in support of public investments in technology-based state capacity (such as biometric payments systems) to reduce corruption and leakages in programmes to help the poor. In many developing countries, the high level of corruption among public officials has affected the performance of these

[1] The exchange rate conversion of US\$/Rupee used is 66.5580. This is the prevailing rate as at 17th October, 2016 sourced from https://www.oanda.com/currency/converter/

[2] According to Google Scholar as at 14th January 2017.

[3] National Rural Employment Guarantee Scheme (NREGS)

[4] Social Security Pensions (SSP)

programmes. According to Niehaus and Sukhtankar (2013), corruption is a marginal cost of social spending and a cause of diversion of substantial sums of money from welfare enhancing programmes such as NREGS. Leakages of funds from its intended use due to corruption in the public system can discourage policy-makers from continuing to finance existing or future programmes (Olken, 2006). Similarly, it is hoped that this replication will provide further insights on the factors that affect the implementation and performance of Smartcards.

II. Brief Description of the Field Experiment

MNS reports the results of a randomized[5] field experiment where Smartcards were rolled out across 8 districts[6] in the Indian state of AP between 2010 and 2012. Across the districts, 296 mandals (sub-districts) were selected from a total of 405 and randomized into treatment (112 mandals, 37.8%), control (45 mandals, 15.2%) and buffer (139 mandals, 47%) groups[7]. The buffer mandals were not included in the analysis, and were chosen so that Smartcards could continue to be distributed beyond the treatment mandals without affecting the control group.

Within each mandal, a fixed number of villages, known as Gram Panchayats (GPs) were selected, producing a total of 880 GPs. Within each GP, 10 households were selected to interview. The households were chosen so that most of households had at least one member who was a recipient of NREGS or SSP benefits. This required selecting households based on state administrative data. It was possible for there to be more than one member of the household who received these benefits, so that the full sample size exceeded 8800. Further, in the course of the survey, some households could not be located. These were identified as "ghost households," possibly because they were non-existent households created so that officials could illegally appropriate funds. Two surveys comprise the data for the analysis. A "Baseline" survey was done before the rollout in 2010. And an "Endline" survey was done two years later. It is important to note that the two surveys, though they surveyed the same GPs, made no attempt to match households across surveys. In the words of MNS, "The resulting dataset is a panel at the village level and a repeated cross-section at the household level" (page 13).

---

[5] The randomization was stratified by districts and socio-economic characteristics of surveyed households

[6] The districts considered are: Adilabad, Ananthapur, Kadapa, Khammam, Kurnool, Nalgonda, Nellore, and Vizianagaram. The 8 districts have a total of 405 mandals.

[7] The key difference between the treatment and control group is the system of payments for NREGS and SSP programs. In the treatment group mandals, payments were made through the "Bank → Technology Service Provider (TSP) → Customer Service Provider (CSP) → Worker" Smartcards enabled channel. The control group payment system channel is from "State → District → Mandal → Gram Panchayat → Worker".

It is also important to note that the rollout was not yet complete by the time of the Endline survey in 2012. Rollout typically occurred at the GP level. Enlistment "camps" were held over a period of 1-2 days. Villagers would attend the camps and get signed up for the Smartcards. The rollout at this level was not random. They were done at the discretion of the GP-level providers. Since providers were compensated based on the number of transactions using Smartcards, we expect that GPs were chosen, at least to some degree, based on the number of potential beneficiaries; and that villagers signed up based, at least to some degree, on the personal benefit that they perceived from using the Smartcards.

In recognition of the non-random nature of the take-up of Smartcards at the GP-level, the "treatment" variable consists of a dummy variable at the mandal level, since the selection of treatment and control mandals was designed to ensure "balance." The corresponding, estimated treatment effect is thus an "intent to treat" (ITT) effect.

The table below summarizes the information above.

Table 1: Sampling Description

	Baseline (2010)	**Endline (2012)**
Districts	**8**	**8**
Mandals	**157**	**157**
• Treatment	112	112
• Control	45	45
Villages (GPs)	**880**	**880**
Households	**10**	**10**
• NREGS	6	6
• SSP	4	4
Sampling Frame	**8,800**	**8,800**
Sample	**8,572**	**8,774**
Unable to survey or confirm existence	1,000	295
Ghost households	102	365
Survey data	**7,425**	**8,114**

III. Method of Analysis and Key Results of MNS

MNS's basic analysis consists of two equations:

(1) $Y = \alpha + \beta Treated + District\ FEs + \gamma PC + error$

and

(2) $Y = \alpha + \beta Treated + District\ FEs + \gamma PC + \delta \bar{Y}^0 + \ error$

where $\bar{Y}^0$ is the GP-level mean of the dependent variable at the time of the baseline and *PC* is the principal component variable that was used to stratify mandals. The key variable here is *Treated*, which takes the value if the individual/household belonged to a mandal that had been selected for treatment.

MNS found that Smartcards had the most pronounced benefits for the NREGS program. While many of the results for the SSP program were not statistically significant, they were generally of the same sign as the NREGS benefits. A summary of their main findings is given below. Only significant findings are reported.

- Reduced time required for beneficiaries to receive payment (NREGS)
SOURCE: Table 2 – individual level data

- Reduced lag time between work performed and payment received (NREGS)
SOURCE: Table 2 – individual-week level data

- Reduced variance in time between work performed and payment received (NREGS)
SOURCE: Table 2 – individual-week level data

- Reduced "leakage" (NREGS)
SOURCE: Table 3 – household level data

- Increased program participation (NREGS)
SOURCE: Table 5 – household level data

IV. Replication Objectives and Research Questions

Our analysis will investigate the following questions:

1) Can we confirm the original results using the data and code provided in the paper?

2) How does the estimated effect differ as one moves from ITT to TOT?

3) Are the results robust?

4) Can we identify moderating factors that contribute to effect heterogeneity?

Question 2 and 3 belong to the category of Measurement and Estimation Analysis (MEA). The last question addresses Theory of Change (TCA). These are described in greater detail below.

IV Pure Replication

4.1 Can we confirm the original results using the data and code provided in the paper?

A detailed list of the main analyses in MNS is given below:

1. Payment logistics, i.e. access to payments
    a. Average time taken to collect a payment (in minutes) [SSP & NREGS]
    b. Average lag (in days) between work done and payment received [NREGS]
    c. Absolute deviation of payment lags (in days) from week-specific median (at mandal's level) [NREGS]
2. Payment amounts and leakages
    a. Amount received by surveyed beneficiaries (in rupees) [SSP & NREGS]
    b. Composite payment leakage (difference between documented official and survey amount received by beneficiaries in rupees) [SSP & NREGS]
    c. Channels of leakages [SSP & NREGS]:
        I. Ghosts == Incidence of ghost households
        II. Over-reporting == Jobcards with zero reported survey payment but positive official payments
        III. Underpayment == Incidence of bribes paid to collect payments
3. Program access
    a. Proportion of households doing NREGS work, i.e. working households
    b. Household willing to work but unable to get NREGS work (in Peak & Slack periods of labour demand) i.e. Involuntary unemployment
    c. Availability of NREGS work to anyone who is willing in the village i.e. Informed employment opportunities among villagers
    d. Did the respondent pay anything to get the NREGS work? i.e. Bribed participation in NREGS
    e. Did the respondents pay anything to receive pension payments? [SSP] i.e. Bribed participation in SSP

The pure replication exercise involves re-estimating the results presented in Table 1-7. The first table presents findings on the usage of Smartcards for NREGS and SSP programs using the official and survey data. Tables 2, 3 and 4 report findings on access to payments (i.e. average time taken to collect a payment using the Smartcards), program benefits (i.e., effect on leakages), and channels of leakage reduction respectively. Table 5 reports findings on access to programs proxied by participation rate in the NREGS and SSP programs. The negative and positive effects of the Smartcards implementation based on opinions of the surveyed participants who are beneficiaries are shown in Table 6. Results of Smartcard implementation by carded vs. non-carded village status are presented in Table 7.

The first question to be addressed in our replication is confirmation of MNS's results. MNS provided a well-organized set of files consisting of data, code, and instructions for reproducing their results. As a result, we are able to exactly reproduce all the tables in their paper.

Table 2: Results of Pure Reproduction Exercise

TABLE/FIGURE	Reproduction Outcome
Table 1	*Reproduced exactly*
Table 2	*Reproduced exactly*
Table 3	*Reproduced exactly*
Table 4	*Reproduced exactly*
Table 5	*Reproduced exactly*
Table 6	*Reproduced exactly*
Table 7	*Reproduced exactly*

V. Measurement and Estimation Analyses (MEA)

5.1 MEA I: How does the estimated effect differ as one moves from ITT to TOT?

For reasons described above, MNS are careful to emphasize that their results are intent-to-treat effects, as the treatment effects are measured at the mandal level. Nevertheless, they do report "non-experimental decomposition" results in Table 7. The estimated "Have SCard, Carded GP" effects can be thought of as estimating the effect of treatment-on-the-treated (TOT), with the caveat that the selection into treatment (Smartcards) was not random. Even though selection into treatment isn't random, these estimates are still of interest because they help to establish "an upper bound" on the benefits of Smartcards. As a result, we plan to explore this further in our replication. We plan to do this two ways.

First, using an analysis identical to MNS, we will estimate the same dimensions of Smartcard benefits reported in Table 7, except that we will compare "Have SCard" to "Does Not Have

Scard" using a different estimation method such as GLMM. These results should be similar to those reported in Table 7. We will then conduct a propensity score matching (PSM) analysis using the same data to measure the estimated effects of Smartcards. The PSM provides an approach, albeit imperfect, for controlling for non-random selection into treatment. It should be useful to help us get a better sense of the "upper bound" of treatment effects associated with the use of Smartcards.

5.2 Are the results robust?

Our replication will analyse the robustness of MNS's results to (i) outliers, (ii) alternative specifications, and (iii) alternative estimation procedures.

5.2.1 MEA II:

Outliers. It is well-known that the presence of a relatively small number of outliers can exert an important influence on estimated effects. Our analysis will employ a number of post-estimation procedures available in Stata for the detection and replacement/removal of outliers, including: (i) discrepancy measures, (ii) leverage measures, and (iii) influence measures (Williams 2016[8]). We will consider both deletion of outliers and winsoring of extreme values.

5.2.2 MEA III:

Alternative Specifications. Household socio-economic characteristics and outcome indicators were used to stratify the randomization of Smartcard payment system across mandals. MNS used the first principal component (PC) of a vector of the mandal characteristics as a regressor to account for differences in outcomes across district. From the descriptive analyses conducted on the baseline household survey data, they reported that there is no significance difference in the mean of the characteristic across treatment and control mandals. This suggests that mandals in the treatment and control groups have the same socio-economic characteristics.

To further investigate this, we propose to test the potential biases that might be created if we exclude the PC variable from the original specification by estimating:

(3) $$Y = \alpha + \beta Treated + \delta \bar{Y}^0 + District\ FEs + error$$

[8] Full description of the techniques can be sourced from https://www3.nd.edu/~rwilliam/stats2/l24.pdf

If the estimated average effect (β) in model (3) differs significantly from (1), it means that the estimated treatment effects are sensitive to the PC indicator. This kind of omitted variable bias could be useful for identifying the existence of heterogeneous treatment effects.

Relatedly, we plan to replace the PC variable with a vector of mandal and GP-related characteristics from the 2001 and 2011 Censuses to estimate the following specification:

(4) $$Y = \alpha + \beta Treated + \delta \bar{Y}^0 + District\ FEs + Mandal/GP\ Characteristics + error$$

A test of joint significance of the mandal/GP characteristics also provides evidence that further investigation into heterogeneous effects is warranted.

5.2.3 MEA IV:

Alternative Estimation Procedures. Modelling quantitative data with clusters (such as districts) using pooled methods often involves accounting for specific group units. It has become a common econometric practice to account for these using either fixed or random effects. There is a plethora of suggestions in literature (e.g., Wooldridge 2010, Greene 2008, Gelman 2005) and some confusing advice (for details see Gelman and Hall, 2007: 245) on the type of group effect to consider.

In analysing randomized cluster dataset where: (i) samples are repeated cross-section[9] (as in MNS[10]), (ii) between and within group correlations are relatively high, (iii) differences in outcomes across group is heterogeneous, and (iv) the response variable follows unknown distribution, the Generalized Linear Mixed Model (GLMM) has been found robust to address these concerns (Stroup, 2012; Gelman & Hall, 2007; Pinheiro & Chao, 2006; Skrondal & Rabe-Hasketh, 2004). GLMM is unique in its modelling of the outcome variable on a set of linear predictors (such as the treatment indicator and socio-economic characteristics of participants in the NREGS and SSP programs), in addition to incorporating both fixed and random effects. When there are no random effects and errors follow the strict classical assumptions (e.g., no between group correlation and constant variance), GLMM is the same as the linear regression model with fixed effects used in MNS.

[9] The dataset employed in the main analysis is not panel. The baseline survey was used to determine the household characteristics at mandal's level. The endline survey was used for all the ITT estimation reported in the study.
[10] See MNS (2016: appendix, page 49) for adopted sampling strategy

For the purpose of this replication study, the sensitivity of the original results will be assessed using the GLMM estimation procedure implemented in the R statistical program. A key motivation for implementing the GLMM procedure is to examine potential bias that might have been created by linear regression with fixed effects when the response variables are correlated across clusters (either at the district or mandal levels).

The modelling of the impact of Smartcards that vary by district can also be achieved using multilevel modelling techniques with hierarchical random effects, such as the Hierarchical Linear Mixed Model (HLMM). GLMM and HLMM are similar in that they are both mixed models (simultaneously incorporating fixed and random effects) and address correlation in the response variable. They differ in their assumption of the distribution of the outcome variable (non-Gaussian or Gaussian), modelling of the response variable to the predictors (link function or directly), variance of the random effects (normal or non-normally distributed), and variation of the intercept/slope (pooled or un-pooled) (for details see Gardiner, Luo & Roman 2009, Gelman & Hall 2007, Schabenberger 2005). In explaining the strengths and weaknesses of the multilevel (hierarchical) model, Gelman (2006:432) notes that HLMM is most applicable for (i) modelling cross-section heterogeneity, (ii) data reduction, (iii) causal inference, and (iv) accounting for systematic unexplained variation among districts.

For the purpose of this replication, we allow our HLMM framework to incorporate heterogeneous intercepts and slopes (treatment effects) as follows from (2) for a two-level hierarchy of measurements (mandals within districts):

(5) $$Y_{ipmd} \sim N\left(\alpha_d + \beta_d Treated_{md} + \kappa X_{md}, \sigma_y^2\right)$$

(6) $$\begin{pmatrix} \alpha_d \\ \beta_d \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_\alpha \\ \mu_\beta \end{pmatrix}, \begin{pmatrix} \sigma_\alpha^2 & \rho\sigma_\alpha\sigma_\beta \\ \rho\sigma_\alpha\sigma_\beta & \sigma_\beta^2 \end{pmatrix}\right)$$

For $d = 1,...,8$, α_d is the intercept for district d and composite for the fixed and random effects; β_d is the slope on treatment indicator for mandal m in district d; κX_{md} is a matrix of covariates ($\gamma\bar{Y}_{pmd}^0 + \lambda PC_{md}$) incorporated in (1) and identical to the predictors used by MNS; and σ_y^2 represents the "within district variation" and σ_α^2 is the variation in "between districts". The districts intercepts and slopes are assumed to be drawn from a normal multivariate distribution that includes a between-group correlation parameter ρ (cf. Equation 6).

VI Theory of Change Analysis (TCA)

6.1 Can we identify moderating factors that contribute to the heterogeneous effects?

Among other things, Theory of Change Analysis (TCA) is concerned with identifying factors that moderate estimated treatment effects. This can be useful for identifying the factors that contribute to treatment success, as well as identifying conditions where the introduction of treatments are most likely to be successful.

MNS (2016, 2910) note that "there was considerable heterogeneity in the extent of Smartcard coverage across the eight study districts, with average rates ranging from 31% in Adilabad to nearly 100% in Nalgonda district". Given the heterogeneous implementation of the biometrically authenticated payment system across districts, there is potential for the impact of the intervention program to differ across districts. Identifying the factors that underlie those differences can be important for the reasons described above.

To investigate the heterogeneous impact of the Smartcard payments intervention programme, we plan to use two different approaches: interaction terms and data decomposition methods. The first approach involves estimating regression models for investigating the effect of outcome variables on treatment indicator with interaction of treatment and district /or mandal indicators as control variable. This is expressed as:

(7) $$Y = \alpha + \beta Treated + District\ FEs + Treated\ *\ District\ Interaction\ Effects + \gamma PC + \delta \bar{Y}^0 +\ error$$

A test of the joint significance of the $Treated * District\ Interaction\ Effects$ will serve as a test of the hypothesis of effect heterogeneity.

An alternative approach is to estimate separate regressions for each of the districts. We plan to decompose the MNS dataset into 8 sub-samples. For each district sub-sample, the original model will be re-estimated using:

(8) $$Y = \alpha + \beta Treated + \gamma PC + \delta \bar{Y}^0 +\ error$$

The estimated β will denote the average treatment effect for a specific district and all the $\beta's$ (i.e. $\beta_1,...\beta_8$) will provide a picture of the heterogeneous impacts of the intervention programme. We will construct confidence intervals for all the districts' treatment effects to inspect and visualize the extent they differ or are close to each other. This analysis is based on the "significance sameness" approach of Hubbard and Lindsay (2013).

The data decomposition strategy may create imbalances in sample sizes across sub-samples and the observations for each dataset might be too small to obtain efficient estimate of the treatment effect for specific district. This concern has the potential of reducing the power of the ITT analysis at district level. It would have been justifiable to conduct power calculation if there is availability of MNS (2016) pilot study dataset. In the absence of such dataset, previous replication studies such as Wood & Dong (2015) employed the survey data to conduct a "post-hoc" power analysis. Although, there are conflicting justification for it use in literature (see Hoenig and Heisey, 2001; and Levine and Ensom, 2001). To avoid the need of conducting a post-hoc power analysis in this replication, we propose to use small-sample size robust estimation method (such as HLMM) as a robustness check.

Our final TCA analysis will build on the results we find in the section on "Alternative Specifications" (cf. Section IVC above). Should the prior analysis identify the importance of mandal/GP characteristics, we will then explore the use of interaction effects to see whether these factors contribute to effect heterogeneity. We will be careful to minimize data mining in our search for moderating effects, and will be careful to provide complete documentation of the steps we follow in our analyses.

VII. Tentative Timeline

The table below presents our planned timeline for carrying out the tasks described above.

Table 3: Planned Timeline

Months (2017)	**Tasks**
April - May	4.1) Can we confirm the original results using the data and code provided in the paper?
May- June	5.1:-MEA I) How does the estimated effect differ as one moves from ITT to TOT?
July - August	5.2:- MEA II - IV) Are the results robust?
September- October	6.1:- TCA) Can we identify moderating factors that contribute to effect heterogeneity?
November - December	Write and modify computer programs used for the replication
November - December	Prepare report and write manuscript

VIII. Conclusion

The objective of this replication is to validate the conclusion in MNS (2016) by using the authors' original data set. Additional analyses will be conducted to assess the robustness and sensitivity of the original published findings. The verification is to establish the credibility of the original findings for policy designs and inform policy formation in other LDCs that are in search of effective anti-poverty programs. The advanced mixed randomized cluster analyses proposed in the replication will provide additional results. The new findings are expected to give clarity and strong justification for the need of building technical state's capacity to promote development interventions, influence policy and provide excellent welfare services at the grass-root.

IX. Disclosure of Interaction with Original Dataset and Code

The investigators of this replication project have accessed the data, code and supplementary materials that the original authors made available on the *American Economic Review* website, https://www.aeaweb.org/articles?id=10.1257/aer.20141346&&from=f

The data and code have only been used for pure replication only.

References

[1]. Carvalho, N., & Rokicki, S. (2015). The impact of India's JSY conditional cash transfer program: A replication study. *3ie Replication Paper*, *6*.

[2]. Clark, T. S., & Linzer, D. A. (2015). Should I use fixed or random effects?. *Political Science Research and Methods*, *3*(02), 399-408.

[3]. Gardiner, J. C., Luo, Z., & Roman, L. A. (2009). Fixed effects, random effects and GEE: what are the differences?. *Statistics in medicine*, *28*(2), 221-239.

[4]. Gelb, A., & Clark, J. (2013). Identification for development: the biometrics revolution. *Center for Global Development Working Paper*, (315).

[5]. Gelman, A. (2005). Analysis of variance—why it is more important than ever. *The Annals of Statistics*, *33*(1), 1-53.

[6]. Gelman, A. (2006). Multilevel (hierarchical) modeling: what it can and cannot do. *Technometrics*, *48*(3), 432-435.

[7]. Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge;New York;: Cambridge University Press.

[8]. Greene, W. H. (2008). *Econometric analysis*, Sixth Edition. Upper Saddle River, NJ: Prentice Hall.

[9]. Hoenig, J. M., & Heisey, D. M. (2001). The abuse of power: the pervasive fallacy of power calculations for data analysis. *The American Statistician*, *55*(1), 19-24.

[10]. Hubbard, R., & Lindsay, R. M. (2013). From significant difference to significant sameness: Proposing a paradigm shift in business research. *Journal of Business Research*, *66*(9), 1377-1388.

[11]. Levine, M., & Ensom, M. H. (2001). Post hoc power analysis: an idea whose time has passed?. *Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy*, *21*(4), 405-409.

[12]. Lewis-Faupel, S., Neggers, Y., Olken, B. A., & Pande, R. (2014). *Can electronic procurement improve infrastructure provision? evidence from public works in india and indonesia* (No. w20344). National Bureau of Economic Research.

[13]. Muralidharan K., Niehaus N. & Sukhtankar S. (2016). Building State Capacity: Evidence from Biometric Smartcards in India. *American Economic Review*, 106(10): 2895-2929.

[14]. Muralidharan, K., Niehaus, P., & Sukhtankar, S. (2014). *Payments infrastructure and the performance of public programs: Evidence from biometric smartcards in india*. National Bureau of Economic Research.

[15]. Niehaus, P., & Sukhtankar, S. (2013). The marginal rate of corruption in public programs: Evidence from India. *Journal of public Economics*, *104*, 52-64.

[16]. Olken, B. A. (2006). Corruption and the Costs of Redistribution: Micro Evidence from Indonesia. *Journal of Public Economics,* 90 (4-5), 853-870.

[17]. Pinheiro, J. C. & Chao, E. C. (2006). *Efficient Laplacian and Adaptive Gaussian Quadrature Algorithms for Multilevel Generalized Linear Mixed Models*. *Journal of Computational and Graphical Statistics, 15*(1), 58-81

[18]. Pritchett, L. (2010). *Is India a flailing state?: detours on the four lane highway to modernization*. Working Paper RWP09-013, Harvard Kennedy School.

[19]. Schabenberger, O. (2005). Introducing the GLIMMIX procedure for generalized linear mixed models. *SUGI 30 Proceedings*, 196-30.

[20]. Skrondal, A. & Rabe-Hesketh, S. (2004). *Generalized Latent Variable Modeling: Multilevel, longitudinal, and structural equation models*. Chapman & Hall/CRC Press.

[21]. Stroup, W. W. (2012). *Generalized linear mixed models: modern concepts, methods and applications*. CRC press.

[22]. Wood, B. D., & Dong, M. (2015). Recalling Extra Data: A Replication Study of Finding Missing Markets. *3ie Replication Paper, 5.*

[23]. Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

[24]. World Bank (2014). *India-Global Financial Inclusion (Global Findex) Database 2014*. http://microdata.worldbank.org/index.php/catalog/2433

[25]. World Bank (2016). *World Development Indicator*. The World Bank Group, Washington D.C.