PREDICTING HOUSEHOLD'S MOBILE BANKING SAVINGS BEHAVIOR IN WESTERN-KENYA: AN ALGORITHMIC APPROACH

Boris Houenou[1] and Ecce F. Djogbenou[2]

Abstract

Tied to household's consumption patterns, saving accumulation drives household's human and business capital investment, and national future growth. From utilitarian to behavioral perspective, digitalization holds promises to unlock saving accumulation in Africa, including for poor. This paper predicts the saving behavior in rural Kenya, taking advantage of machine learning algorithms to examine a massive set of features known theoretically and empirically to affect household's saving behavior. The algorithm fits a generalized regression model through a penalized maximum likelihood. The work anchors in economic theory on saving but differs from the method standpoint to elicit mobile banking saving take-up and saving accumulation.

Data covered thirty-six villages in Western Kenya. Predictors clustered into livelihoods, assets inventory, formation and use of assets, income generation and its use, food, composition, intake and nutrition, the housing quality and tenureship, water and sanitation, energy use, family structure, social status architecture, adverse shocks to agricultural production (crop and livestock loss events) and demographics. Data include 1600 households from six different communities and 7,700 quarterly observations from 2013 to 2015.

Although the results point to shreds of evidence that corroborates previous findings, we highlight in this paper, mobile banking's role in the saving accounts ownership, and saving accumulation, as well as the effects of age and employment status. Mobile phone ownership is a strong and significant predictor of saving account ownership. However, the mobile phone is a vehicle to fluctuating saving balance and encourages dissaving. Moreover, being young and in the workforce increases the mobile banking saving take-up but also the dissaving behavior, compared to older age group. This work subsets features that could help improve product or policy design towards a better financial inclusion of rural poor.

JEL Codes C4, C52, E21

Keywords: Digital and Mobile Banking-Household's Saving-Elastic Net-LASSO

[1] *School of Economic Sciences, Washington State University, and ViaSat Inc,*
Email: borishouenou@gmail.com.
[2] *Ecole Nationale d'Administration Publique, Quebec, Canada*

Highlights

Owning a mobile phone is the single most important predictor of household saving behavior through mobile banking.

Owning a mobile phone increases the likelihood to own a saving account.

Having a mobile phone leads to dissaving by increasing the frequency and amount taken out of the saving account.

Fewer elderly household's heads tend to own a saving account. When they do, they tend to have a higher balance than the younger household's heads.

Introduction

Besides individual welfare, saving accumulation purportedly drives economic growth. At a household or individual level, saving is tied to consumption patterns, including smoothing but also determines productive investments in human and business capital (Karlan et al., 2014). At the same time, future economic growth depends on the household's saving (Karlan et al., 2014). If market failure, transaction costs, risk perception and lack of trust or regulatory barriers have excluded households in developing countries from traditional saving products (Demirguc-Kunt and Klapper, 2012), recent digitalization has unlocked mobile banking, including to the rural households. Research and practical interests are to understand what drive the take-up of saving accounts using mobile banking, how substantive is the saving accumulation through those products and whether the preferences and patterns vary depending on the age, gender or being in the workforce. Through utilitarian lens or incentives and biases driven approaches, these questions have attracted prolix research.

Two main threads of literature take prominence in the works on savings: the classical utilitarian and behavioral approach. While the utilitarian approach, mostly built on the intertemporal utility is pervasive in macroeconomics, behavioral economics concerns with household's incentives and biases to save or undersave. Dominantly, the theory of the intertemporal utility model concerns with precautionary saving and liquidity constraints (Hall and Mishkin,1982; Skinner 1988; Zeldes, 1989; Caballero 1990). Also, bequest and habits formation and the life cycle hypothesis (Modigliani, 1966) are other subtopics of that theory. Bequests, precautionary savings and liquidity constraints, permanent income and habit formation models do not fulfil the expectation of explaining the peculiar traits of saving observed in developing countries households (Deaton, 1989). In developing countries, households are larger and are more likely to span several generations, retirement saving, or intergenerational transfers are less evident or uncertain and cyclical income leads to higher precaution in consumption smoothing. Above all, households are more liquidity constrained, and borrowing constraints are severely binding.

To explain saving patterns through incentives and biases, behavioral economics challenges the exponential discounting in the intertemporal utility model. Individual time inconsistency suggests non-exponential discounting. Many experimental studies indicate that the individual has preferences that reverse as the date of decision-making nears (Ashraf et al., 2003). Thereafter, the hyperbolic discounting and mental accounting models emerge in the literature (Lowenstein and Thaler, 1989; Thaler, 1992, 1990; Laibson, 1997; O'Donahue and Rabin, 1999).

The framework has been modified to accommodate data as much as possible. Also, developing countries have entered the digital world, and many predictions may or not hold about their households' saving behavior. Some recent experimental studies measure the effects of saving devices on the saving for preventive healthcare (Dupas and Robinson, 2013a). They found -using data from a field experiment in Kenya -that with simple, informal saving technologies individuals increase their saving substantially to invest in preventive healthcare and reduce vulnerability to health shocks. The impact is increased when the experiment combines or uses more "sophisticated" and group base savings and credit schemes. Their investigation is at the nexus of some essential thoughts from a utilitarian standpoint that we explore above. Life cycle, precautionary, and potentially habit formation are all possible long-run effects that could be delineated from the outcomes of their experiments.

In the core of the recent literature on saving through mobile banking are the determinants of saving products take-up, dissaving, as well as the effects of gender, age, and employment status on saving patterns. Overall, mobile banking saving has shown some impact on welfare (Asongu and Nwachukwu, 2016; Steinert et al., 2018), enhances pro-saving attitudes (Steinert et al., 2018) and improves consumption smoothing (Batista and Vicente, 2020). These effects are enhanced in the presence of financial literacy and some education level (Zins and Weil, 2016), especially for poor and low-income groups. A growing body of evidence suggests that perceived risk and lack of trust reduce mobile banking saving products' take-up (Nyoka, 2018; Kibicho and Mungai, 2019; Seck Fall et al., 2020). Although the perceived ease of use tends to increase the adoption of saving products, those studies have pointed to substantial perceived risk and lack of trust either because of weak financial literacy or information gaps.

Saving through mobile banking decreases the propensity to use informal saving mechanisms (Mbiti and Weil, 2016). It also increases the propensity of disadvantaged groups such as rural, female, less educated individuals, and individuals with irregular income to save for health emergencies (Ky et al., 2018). However, it does not lead to saving accumulation (Mbiti and Weil, 2016; Steinert et al., 2018; Nyoka, 2018; Batista and Vicente, 2020), because of the frequency and convenience of the use of the mobile phones (Ouma et al., 2017). Education mediates the gender effects of the mobile banking on saving, and the adoption rate of the products by rural women is lower than men, even though when female adopt saving products, they use it more (Seck Fall et al., 2020). Regarding age and employment status effects, mobile banking saving take-up rate is higher in a younger population (Karakara and Osabuohein, 2018) compared to older age category. However, provided that an individual has a mobile banking saving account, being out of the workforce leads to more saving accumulation.

Notwithstanding that many models' predictions found solid ground in explaining saving behavior at micro-level, this paper has a thread of contributions. It extends the literature in understanding mobile banking saving behavior by harnessing regularized regression in machine learning. The method approaches bias reduction in estimation from a different angle than precedent works. The algorithmic approach does not concern too much with the data generation process. We study mobile banking saving take-up, saving accumulation, as well as age and gender effects on the two variables using a high dimension design matrix to predict saving behavior of households in Western Kenya optimally.

Using unique survey data from rural Kenya, we explore how much the features associated with various saving models in the literature (and usually tested individually) predict saving behavior in rural households in developing countries. We explore the decision to save and saving accumulation through mobile banking, leveraging the elastic net penalized algorithm. The algorithm selects the features and operates the regularization of the model such that the holdout results are close enough to the training sets outcomes. From a method standpoint, our paper's use of regularized regressions algorithms to study mobile banking saving often approached in with classical econometrics tools is novel. Features selection helps us to subset the main variables that predict the saving behavior in western Kenya. Harnessing the similarity of the selected variables with the covariates used in studies that test different saving hypotheses our work could implicitly inform which model has the highest likelihood in predicting households' saving behavior in developing countries. As policy implications, this paper subsets features to consider in designing saving products leveraging digitalization that could benefit financial institutions[1] operating in pensions plan, medical savings or alike.

The rest of the paper is organized as follows. In section 1, we survey the main saving behavior models by briefly introducing the theories and a selection of empirical results. In

---

[1] Who usually implicitly or explicitly incorporate the saving behavior in the products development and modelling (Ashraf et al. 2003)

section 2, we briefly expose the regularized regression methods and their versions. Section 3 describes the data. Section 4 exposes the results of the saving behavior, notably, mobile saving take-up, saving accumulation as well as age effects, using the optimal model. Section 5 gives the concluding remarks.

1. **Savings Behavior Survey**

We start with the traditional intertemporal utility model. The modifications of that model, such as the certainty equivalence model and Deaton's inputs, are discussed. We then examine the macro-level implications through the lens of life cycle saving behavior.

1.1. Intertemporal Utility Model and Permanent Income Hypothesis

A household solves an intertemporal utility maximization problem choosing consumption and assets levels. The model relies on three main assumptions: (1) additively separable utility function, i.e. that consumption is enjoyed when it takes place, independent of the history of consumption, discounting tomorrow's consumption by a factor-beta that measures the impatience or time preference and represents a bias against savings, (2) borrowing is possible at the market time preference *r* to assure smoothing consumption as long as the capital market is perfect, and (3) the agent is risk-neutral (in case of the quadratic utility function in consumption), or uncertainty is included.

Formally, the agent's optimization problem is written as:

$$\underset{C_t, A_t}{\mathrm{Max}}\ E_t[U(C_t, C_{t+1} \dots C_T; B|Z_t, Z_{t+1} \dots Z_T)]$$

subject to: (1)

$$C_{t+s} + A_{t+s+1} = Y_{t+s} + (1 + r_{t+s})\ A_{t+s}$$

where:

C_{t+s}: consumption at time $t+s$

Y_{t+s}: income at time $t+s$

A_{t+s}: assets held at the beginning of period $t+s$

Z_{t+s}: variables affecting utility of consumption such as household demographics

B: bequest left to younger generations, $B \geq 0$

r_{t+s}: interest rate at time $t+s$

$E_{t:}$ statistical expectation operator taken at time t

The solution to the problem yields Euler equation which represents the arbitrage between today and tomorrow consumption. This formulation is the workhorse of saving behavior analysis. The alteration of the assumptions gives rise to other variants of the model.

1.2. Certainty Equivalence Model (CEQ), Liquidity Constraints and Precautionary Savings, and Habits Formation

When the utility function is quadratic in consumption and the equivalence scale is introduced in the demographics of the agent, then we have the primary form of certainty equivalence model. The model is attributed to Hall (1978) and Flavin (1981). The utility function becomes $U(C_t, Z_t) = U(C_t/\alpha(Z_t))$. The equivalence scale allows the study of the behavior for an agent with or without children[2], marital status but also the bequeathing behavior. The power of the certainty equivalence model resides in its simplicity, notably when the utility function is quadratic in the consumption. For instance, if the demographics are time-invariant or if we can guarantee that they are stable, only agent's lifetime resources, and time preference determine the path of optimal consumption. Permanent income carries more weight than current income. Also, the optimal consumption follows a random walk as long as the subjective individual time

[2] For instance, an adult is 1 equivalence scale.

preference is the market's time preference such that $\beta_t = 1/(1 + r_t)$. The model while appealing for many empirical tests reveals that demographics are the backbone of the power of the empirical tests and it is hard to know to which extent demographics and which demographics fit the data very well. Moreover, the risk neutrality is so simplistic that data do not support them that well.

The precautionary savings and liquidity constraints model relax the risk type assumptions. It predicts different savings behavior for low and high lifetime income individuals. Precautionary savings posit that risk-averse individuals suffer a more significant utility decline, from cutting consumption, than they obtain utility increase (from comparable size increase in consumption). That entails that agents have a preference to hold assets (or borrow less) and have an increase in consumption over time rather than smoothing the consumption. The liquidity constraints also imply that agents consume less today. In both cases, the agents consume less now than they would like to either because they cannot borrow (liquidity constraints) or are scared of borrowing out of concern that they will not be able to pay back in the future without having a substantial decline in consumption (Coleman, 1998).

The CEQ model -though is not widely accepted (Browning and Lusardi, 1996) -draws from the life cycle saving model and might not be convincingly associated with the permanent income hypothesis (PIH) argument. In fact, Friedman (1957) allows for liquidity constraints and not for precaution motive of saving (Browning and Lusardi, 1996). The extension to the bequest analysis will not be exposed here since the structure of saving behavior in developing countries differ from the one in developed countries where accumulation, bequests and retirement motives make the bulk of the saving. We expose the alteration of the standard intertemporal model by Deaton (1989) and the recent field experiments' results in testing and measuring saving behavior in developing countries.

Built on the main differences drawn from the classical literature in developed countries, Deaton's idea of saving behavior in developing countries' households embeds borrowing constraints -with the particularity that households cannot borrow and do not have access to infinite credit -and a time preference that is higher than the market interest rate which is a necessary requirement for the borrowing constraint to affect the solution (Bewley, 1977). His model attempts then to tell a different story than the one of the permanent income even in developing countries.

The optimization problem of the household is to choose an intertemporal consumption as follows:

$$u = E_t[\sum_t^{\infty}(1+\delta)^{-t}v(c_t)]\} \tag{2}$$

$subject\ to$

$$A_{t+1} = (1+r)(A_t + y_t - c_t) \tag{3}$$

$$A_t \geq 0 \tag{4}$$

Equation (2) is the intertemporal expected utility of the household where E_t is the expectation operator, u is the utility, δ is the rate of time preference, c_t is consumption, v is the instantaneous utility associated with consumption c_t, and t the time index.

Equation (3) is the classical resources constraint where A_{t+1} is the wealth at time $t+1$, r is the market time preference, y_t the income at time t and A_t is the wealth at time t.

Equation (4) is the constraint that the wealth at any given time t is positive.

This setting implies two cases that Deaton (in his own words) summarizes as: (1) *"In each period, the consumer would like to borrow but cannot; even if all wealth and current income are consumed, the marginal utility of an additional unit of current consumption is greater than the expected marginal utility to be derived by saving that rupee until tomorrow".*

Therefore, only the sum of assets and income make the total consumption, and the saving is zero or negative, no assets are carried forward, and marginal utility is not equated across periods. *(2) "The consumer does not want to borrow; consumption is less than total cash on hand... Of course, this does not mean that the borrowing constraints have no effect. The expected marginal utility of future consumption is different from the unconstrained case because the future contains the possibility of being unable to borrow when it would be desirable to do so."*

Formally the two cases correspond to the two Euler equations below (equations 5 and 7):

$$\lambda(A_t, y_t) \geq E_t[\frac{(1+r)\lambda(c_{t+1})}{(1+\delta)}] \tag{5}$$

$$c_t = A_t + y_t;\ s_t = y_t - c_t \leq 0;\ A_{t+1} = 0 \tag{6}$$

$$\lambda(c_t) = E_t[\frac{(1+r)\lambda(c_{t+1})}{(1+\delta)}] \tag{7}$$

where λ is the marginal utility of the consumption or alternatively the price of value of consumption and s_t is saving. Simplifying things, the reality of the household's saving behavior in a developing country boils down to:

$$\lambda(c_t) = Max\{\lambda(A_t, y_t), E_t\left[\frac{(1+r)\lambda(c_{t+1})}{(1+\delta)}\right]\} \tag{8}$$

If Deaton's (1989) model finds multiple empirical shreds of evidence in developing countries, the emerging literature on savings in developing countries tends to stress the role played by households' behavior on saving. Using experimental studies, random experimental design, that stream of studies proves that household behavior towards saving, saving product take-up and the impact of saving are altered (Karlan et al., 2014)

1.3. Mobile banking saving seen as incentives and bias driven

Although poor and rural households have a surplus (Banerjee and Duflo, 2007) that could find its way into saving, many barriers play as constraints to saving accumulation in unprivileged and vulnerable groups (Karlan et al., 2014). Predominantly, constraints are a supply-side question that runs from transactions costs attached to the conventional banking systems, either pecuniary (Dupas *et al*. 2012; Dupas and Robinson, 2013b ; Prina, 2013) or non-pecuniary such as proximity of the services and convenience (Schaner, 2013). When fees and costs associated with the savings are either reduced, subsidized or removed, savings patterns are sensitively affected, suggesting individual' response to incentives.

At the surge of the digitalization and the consequential rise of mobile banking saving, a clear path has emerged to uphaul the transactions costs and, then, increase accumulation and take-up of saving products. However, lack of trust plays as moral hazard, this time as a demand-side problem, to moderate the expectations of saving revolution (Brune *et al*. 2013; Nyoka, 2018; Kibicho and Mungai, 2019; Seck Fall et al., 2020. Also, regulatory barriers and social constraints both intra and inter-households' distribution of bargaining power (Jakiela and Ozier, 2012; Dupas and Robinson, 2013a; Ky, 2019) are at play in the saving behavior. By the attached convenience and ease of use (Ouma et al., 2017; Kibicho and Mungai, 2019; Seck Fall et al., 2020), mobile banking saving exhibit some promise to fulfil the saving revolution for the poor.

Nevertheless, even when the demand and supply barriers are reduced, saving is prone either to biases and psychological costs or arbitrage. Among the biases that affect the mobile banking saving products take-up and saving accumulation are the preferences (present bias notably), expectations bias, price perceptions and inattention to savings, purely related to decision-making (Karlan et al., 2014). When information gaps and financial literacy are improved (Fernandes et al. 2013; Zins and Weill, 2016; Seck Fall et al., 2020) it is a legitimate expectation that biases could be lessened, especially price perception and inattention to savings biases. As a tool that democratizes saving adoption, mobile banking holds promises that need better characterization to be unlocked.

2. **Elastic Net as Regularization and Features Selection Method**

Elastic net is a flexible regularization and features selection method that combines both L1 and L2 norm penalizations (Zhou and Hastie, 2005). We use, in the first instance, L1 and L2 penalizations, and then, the elastic net regularization which is the linear combination of the two latter. The continuous penalization using L2, a well-known approach in economics,

is the ridge regression (Hoerl & Kennard, 1988). It tends to correct for the flaws of ordinary least square (OLS) to achieve a better prediction through a bias-variance trade-off (Zhou and Hastie, 2005). By allowing a small bias in the cost function as penalty, the ridge regression reduces overfitting allowing the model to learn better the pattern of data instead of the noise. The penalty, calculated as lambda times the square of the slope, controls the severity of the penalization. In the absence of such a penalty, especially when one has a high dimension design matrix prone to instability, the line of best fit has a steeper slope and is non-robust to a small change in features matrix.

Although, the features matrix becomes stable one cannot guarantee a parsimonious model, since ridge regression tends to keep all the predictors in the model. Some other approaches, like the best subset to variables selection in modeling, involve discreteness that produces sparse models. Because the ridge regression has historical use in economics and is a subset of the elastic net method, we do not discard this algorithm in our estimation suite.

The least absolute shrinkage and selection operator (LASSO) is a more conservative, and performant approach than ridge regression in some circumstances (Tibshirani, 1996). The LASSO estimator -which estimates the linear regression coefficients with penalty calculated as the absolute value of the slope -often yields solutions with some coefficients being precisely zero. To that unsatisfactory outcome, Zhou and Hastie (2005) outline three more drawbacks of the LASSO: (1) the LASSO cannot select more predictors than the sample size, because of the nature of the convex optimization problem; (2) the grouping effect in the LASSO is null, *i.e.* when there is a group structure among the predictors (say correlated variables for example), the LASSO estimator usually selects only one predictor from a group of correlated predictors; and (3) when the predictors are highly correlated, the LASSO estimator performs unsatisfactorily. Note, though that in our paper, the number of observations (*n*) is greater than the number of predictors (*p*).

Altogether, ridge regression and the LASSO perform poorly when the number of predictors is higher than the number of observations. Luckily, in our paper, the number of predictors is smaller than the size of the sample. The elastic net algorithm which combines both methods tends to give better results. Because of the dimensions of our design matrix, not only we can satisfactorily run LASSO and ridge regression, but also improve the accuracy of our prediction using elastic net.

The elastic net penalty is the convex combination of ridge regression and LASSO penalties. Our empirical predictive model is as follows:

$$\widehat{\mathbf{Sr_i}} = \widehat{\boldsymbol{\beta}}_{\mathbf{0}} + \sum_{\mathbf{j=1}}^{\mathbf{J}} \widehat{\boldsymbol{\beta}}_{\mathbf{j}} \mathbf{X_i} + \sum_{\mathbf{k=1}}^{\mathbf{K}} \widehat{\boldsymbol{\beta}}_{\mathbf{j}} \mathbf{Y_i} + \sum_{\mathbf{l=1}}^{\mathbf{L}} \widehat{\boldsymbol{\beta}}_{\mathbf{j}} \mathbf{Z_i} + \sum_{\mathbf{m=1}}^{\mathbf{M}} \widehat{\boldsymbol{\beta}}_{\mathbf{m}} \mathbf{Q_i} + \sum_{\mathbf{n=1}}^{\mathbf{N}} \widehat{\boldsymbol{\beta}}_{\mathbf{m}} \mathbf{R_i} \tag{9}$$

where:

$\mathbf{X_i}$: demographics and household' descriptive variables

$\mathbf{Y_i}$: shock variables, namely natural hazards or circumstances that have generated loss in assets or production

$\mathbf{Z_i}$: income and earning variables for household i

$\mathbf{Q_i}$: consumption and expenditures variables for household i

$\mathbf{R_i}$: commitments devices or saving tools. This study uses mobile phones which we know is needed for adopting and using Mpesa mobile banking which covers more than 60% of the households in the sample.

Sr: household member saving account ownership or balance.

Let $\widehat{\boldsymbol{\beta}}$ represents the vector of the coefficients:

The ridge regression estimator solves for $\widehat{\boldsymbol{\beta}}$ such that:

$$\widehat{\boldsymbol{\beta}} = \underset{\boldsymbol{\beta}}{\mathbf{argmin}}(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})^{\mathbf{T}} - \boldsymbol{\lambda}_{\mathbf{1}}|\boldsymbol{\beta}|^{\mathbf{2}} \tag{10}$$

where: $|\beta|^2 = \sum_{i=1}^{p} \beta_i^2$

The LASSO estimator is such that:

$$\widehat{\boldsymbol{\beta}} = \underset{\boldsymbol{\beta}}{\mathbf{argmin}}(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})^{\mathbf{T}} - \boldsymbol{\lambda}_{\mathbf{2}}|\boldsymbol{\beta}|_{\mathbf{1}} \tag{11}$$

where $|\boldsymbol{\beta}|_{\mathbf{1}} = \sum_{\mathbf{i=1}}^{\mathbf{p}} |\boldsymbol{\beta}_{\mathbf{i}}|$

The Elastic net estimator is such that:

$$\widehat{\boldsymbol{\beta}} = \underset{\boldsymbol{\beta}}{\mathbf{argmin}}(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})^{\mathbf{T}} - \boldsymbol{\lambda}_{\mathbf{1}}|\boldsymbol{\beta}|^{2} - \boldsymbol{\lambda}_{\mathbf{2}}|\boldsymbol{\beta}|_{\mathbf{1}} \tag{12}$$

If we let $\alpha = \frac{\lambda_2}{\lambda_1 + \lambda_2}$, then we can rewrite (12) as:

$$\widehat{\boldsymbol{\beta}} = \underset{\boldsymbol{\beta}}{\mathbf{argmin}}(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})^{\mathbf{T}} \text{ subject to } \boldsymbol{\alpha}|\boldsymbol{\beta}|^{\mathbf{2}} + (\mathbf{1} - \boldsymbol{\alpha})|\boldsymbol{\beta}|_{\mathbf{1}} \leq \mathbf{t}, \text{for some } \mathbf{t}. \tag{13}$$

Equation (13) indicates that the Elastic net penalty is a convex combination of the penalties of the ridge regression and the LASSO problems.

As it appears, we recover a ridge regression problem when $\alpha = 1$ and the LASSO problem when $\alpha = 0$. In this paper, we consider those cases as well as $\alpha = 0.5$ for different tuning (different lambda) to select the most powerful predictors of the saving behavior in

Western Kenya. The design matrix is a 7,700 by 117 i.e. we have n= 7,700 observations and p = 117 parameters.

We use the glmnet package suite in R which is designed to run algorithms for elastic-net regularized generalized linear models (Friedman et al., 2010; Simon et al., 2011; Tibshirani et al., 2012; Simon et al., 2013). The algorithm fits a generalized regression model through a penalized maximum likelihood. Not only, it is the most comprehensive algorithm suite, but it can handle all shapes of data, including exceptionally large and sparse data matrices as we have in this situation. The optimization search uses cyclical coordinate descent in a path-wise fashion. Mobile banking saving account ownership was run under a binomial distribution to training the model while the saving accumulation was run under Gaussian distribution. While mobile saving account ownership was run as a classification problem, the saving accumulation was predicted as a regression problem.

The glmnet algorithm could be sensitive to the scale of variables. Therefore, we have standardized the design matrix before fitting the modeling sequence. Note that the coefficients are systematically transformed back to their scale in the input data. We hyper-parametrize the algorithm to allow an intercept for each model and to limit the number of lambdas to try to 100. We have not supplied a sequence of lambdas, leaving the algorithm to run on the default sequence, because this approach is less computationally expensive and guarantees a fast run of the algorithm. The convergence threshold value of the algorithm is 1e-07. If the maximum change in the objective function after a coefficient update is less than the convergence threshold value times the null deviance, the inner coordinate-descent loop will continue. But the number of iterations was tuned to 100,000. For the rest of parameters of the algorithm, we use default values. The numerical optimization on the gradient of the cost function uses the Newton algorithm, which requires the full Hessian matrix to perform. Finally, we use cross-validation as a robustness check to our estimation by leveraging out-samples (holdout) data.

3. Data Description

The data come from two essential sources: The Human Population-Based Infectious Disease Surveillance (PBIDS) and the Socio-Economic Survey (SES). The PBIDS and the SES are linked at the household level and discussed further in Thumbi et al., (2015). Formed by a conjoint effort of US center of Disease control and Prevention (CDC) and the Kenya Medical Research Institute (KEMRI), the PBIDS initiative is a morbidity surveillance survey that has recorded health events, treatment, treatment-seeking behavior of households and healthcare costs of six different communities in Western Kenya. It provides, for our study contingent situations and shock events profile of the households, allowing us to study saving for unforeseen circumstances as well as shocks to assets that could affect the saving behavior.

The SES built on the PIBDS to provide a social, economic perspective. It covers most aspects intersecting the main drivers that the literature on household's saving behavior has considered. Data include saving products take-up, use and saving accumulation, livelihoods of households, inventory of assets, formation and use of assets, income generation and use, food intake, food composition and nutrition, main living assets inventory especially livestock, housing quality and tenureship, water and sanitation, energy use, family structure, social status, special events like shocks to preferences (crop and livestock loss events) other demographics. In total, the data include 1600 households from six different communities and 7,700 quarterly observations covering years 2013, 2014 and 2015. Because of the age filter we use in this paper, retaining only individuals at least aged to be part of the workforce, the final sample is about 4000 quarterly observations.

We merge the two datasets at household ID and quarter levels. We, later, transform some variables to obtain the variables we use in our analysis. For instance, we compute on-farm cash income, permanent income as lagged household income, macro-nutrients intake, food equivalent scale, expenditures, and saving balance. A total of 123 variables[3] are used in the predictive models.

3.2. Descriptive Statistics of Saving Behavior Data

We provide in figures 1 and 2 the distribution of saving account ownership and the saving account balance. The distribution follows the theme of our analysis which considers the overall sample (16+ years old), individuals comprised in the workforce (16-65 years old) and the senior

[3] Member1 Age, Member1 Gender, Member1 Education Level, Member1 Marital Status, Member1 Family Role, Member1 Village Role, Member1 Primary Occupation, Member2 Age, Member2 Gender, Member2 Education Level, Member2 Marital Status, Member2 Family Role, Member2 Village Role, Member2 Primary Occupation, Member3 Age, Member3 Gender, Member3 Education Level, Member3 Marital Status, Member3 Family Role, Member3 Village Role, Member3 Primary Occupation, Member4 Age, Member4 Gender, Member4 Education Level, Member4 Marital Status, Member4 Family Role, Member4 Village Role, Member4 Primary Occupation, Member5 Age, Member5 Gender, Member5 Education Level, Member5 Marital Status, Member5 Family Role, Member5 Village Role, Member5 Primary Occupation, Total number of male children (<5 yrs) in the household, Total number of female children (< 5 yrs) in the household, Total number of male children (5-10 yrs) in the household, Total number of female children (5- 10 yrs) in the household, Does the household currently own Plough or other Implements?, Number of Units in Usable Condition of Ploughs and other Implements, Market value would you expect if bought or sold these goods today, Does the household currently own Bicycles?, Number of Units in Usable Condition of Bicycles, Market value would you expect if bought or sold these goods today, Does the household currently own Vehicle(s)?, Number of Units in Usable Condition of Vehicles, Market value would you expect if bought or sold these goods today, Does the household currently own Radio(s)?, Number of Units in Usable Condition of Radio(s), Market value would you expect if bought or sold these goods today, Does the household currently own Tractor(s)?, Number of Units in Usable Condition of Tractors, Market value would you expect if bought or sold these goods today, Does the household currently own Mobile Phone(s)?, Does the household currently own Mobile Phone(s)?, Number of Units in Usable Condition of Mobile Phones, Market value would you expect if bought or sold these goods today, Does the household currently own Motor Cycle(s)?, Number of Units in Usable Condition of Motor Cycles, Market value would you expect if bought or sold these goods today, Market value would you expect if bought or sold these goods today, Market value would you expect if bought or sold these goods today, Does the household currently own Television(s)?, Number of Units in Usable Condition of Televisions, Number of Units in Usable Condition of Televisions, Market value would you expect if bought or sold these goods today, Does the household currently own Television(s)?, Number of Units in Usable Condition of Televisions, Market value would you expect if bought or sold these goods today, Number of mud wall grass thatched buildings, Number of Mud Walls and Iron Roof Buildings, Number of mud wall plastered iron roof buildings, Number of Stone, Brick or Concrete Walls Buildings, Number of other types of buildings, Latrine, Electricity, Primary Drinking Water Source, Member1 Off-Farm Net Income over the last 3 Months, Member1 Time Spent per Week earning this income, Member1 Time Spent per Week Working off the Farm in Trade for other Goods and Services, Member2 Off-Farm Net Income over the last 3 Months, Member2 Time Spent per Week earning this income, Member2 Time Spent per Week Working off the Farm in Trade for other Goods and Services, Member3 Off-Farm Net Income over the last 3 Months, Member3 Time Spent per Week earning this income, Member3 Time Spent per Week Working off the Farm in Trade for other Goods and Services, Total from Children Off-Farm Net Income over the last 3 Months, Total from Children: Time Spent per Week for Cash Income off the Farm, Total from Children: Time Spent per Week Working off the Farm in Trade for other Goods and Services, Household Total Off-Farm Net Income over the last 3 Months, Household Total Time Spent per Week for Cash Income off the Farm, Household Total Time Spent per Week Working off the Farm in Trade for other Goods and Services, Does any Household Member Maintain a Savings Account/MPesa/Airtel Money or any mobile money transfer account?, What is the Current Total Household Saving Balance?, Difference between cash income and the total expenditures, Has any Household Member obtained a Loan in the Last 3 Months?, Why was the Loan taken?, During the last 3 Months, what is the estimated value lost in household goods due to flood, fire, theft, other?, Cooking Fuel Cost, Clothes Cost, Health Care Cost, Education Cost, Other Household Cost, Total Food Expenditures, Total Calves owned by the household, Total Heifers owned by the household, Total Bullocks owned by the household, Total Cattle owned by the household, Total Goats owned by the household, Total Sheep owned by the household, Total Poultry owned by the household, Total Donkeys owned by the household, Total Other owned by the household, How Many Acres did you Plant in Crops for the Household During the Most Recent Crop Production Season?, How many of these Acres are Owned by You/Your Household?, Is Part of this Land Rented?, How many of these Acres are Rented?, Have you Lost any Household Crops Due to Drought, Floods, Wild Animals, Fire or any other thing in the Last 3 Months?, Estimated Value of Loss.

individuals (65+ years old) assumed to be out of the workforce. The idea is to tease out any life cycle difference in the saving behavior in the population.

Overall, senior individuals have more in saving account than individuals in the workforce or the pooled sample. For instance, about 55% of individuals 65+ years old have $100 in saving balance for only 20% in the workforce (16-65 years old) or in the overall sample (16+ years old). Similarly, 30% of older have no saving balance against 58% in the workforce and 55% in the overall sample.

Regarding the saving account ownership, about 55% of an individual of 65+ have a saving account in comparison with 50% for people in the workforce and 50% of the overall sample. Table 1 reports descriptive statistics on age, household size and quarterly net off-farm income. The median age of the head of the household is 54 years old, which means that 50% of the sample is still in the workforce. The average size of the household is 5.23, while the median size is 5 individuals. A household's quarterly off-farm net income is about 10,900 Ksh, an equivalent of $100.

Figure 1: Distribution of Saving Account Ownership

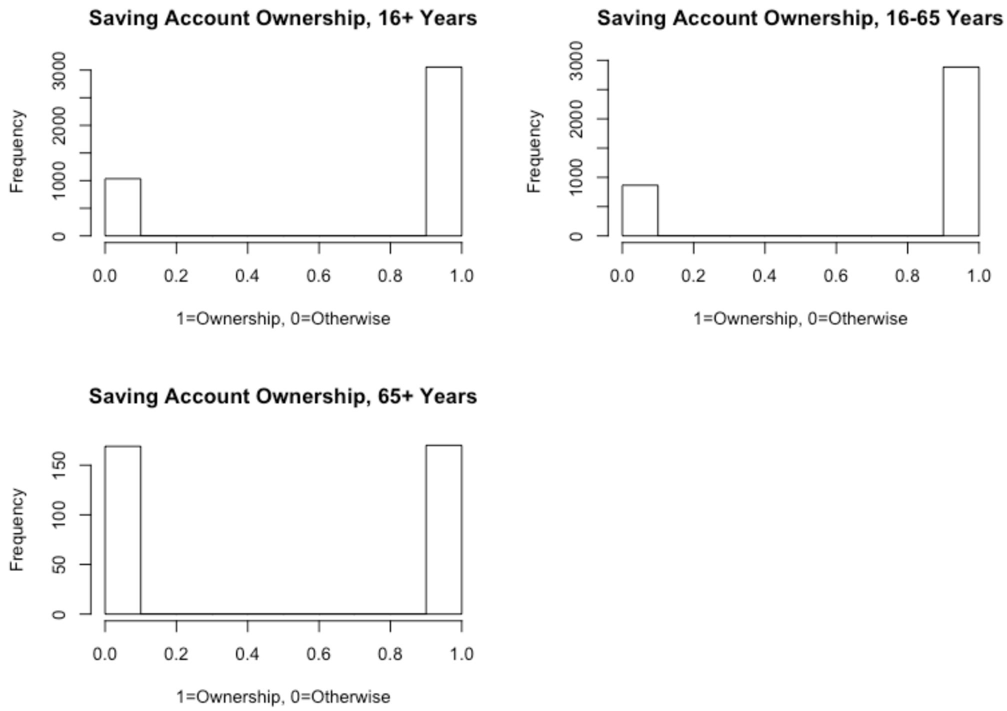


Figure 2: Distribution of Saving Balance

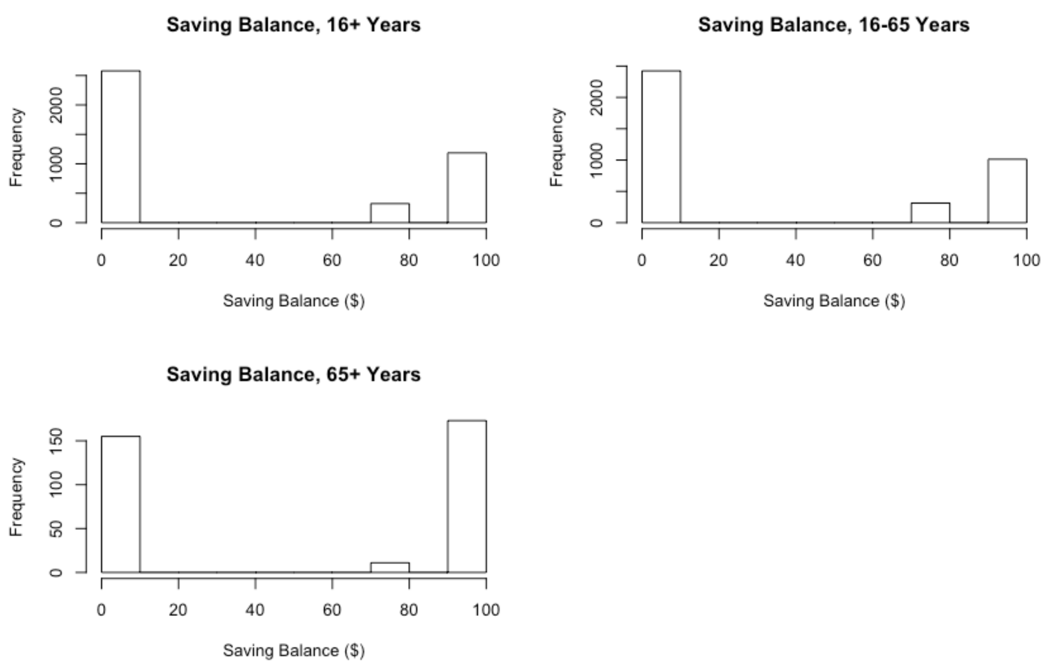


Table 1: Descriptive Statistics

	Overall (N=4101)
Age Head of the Household	
Mean (SD)	53.9 (14.1)
Median [Min, Max]	54.0 [14.0, 89.0]
Missing	2 (0.0%)
Size of the Household	
Mean (SD)	5.23 (2.35)
Median [Min, Max]	5.00 [1.00, 19.0]
Missing	2 (0.0%)
Household's Off Farm Net Income (Ksh)	
Mean (SD)	10900 (23400)
Median [Min, Max]	0 [0, 165000]
Missing	12 (0.3%)

4. Results

Mobile phone ownership is a powerful predictor of household's saving behavior in Western Kenya. Among the predictors, the most regularized model predicts that ownership of a mobile phone is the most important feature for the household's saving behavior. However, the direction of the prediction reverses for saving accumulation. While mobile phone ownership positively predicts the saving account ownership, having a mobile phone leads to a reduction of the saving balance.

4.1. Saving Account Ownership Prediction

Figures 3-5 plot the most important selected features which have the highest prediction weight of the saving account ownership.

Figure 3: Top Predictors of Saving Account Ownership in Overall Sample

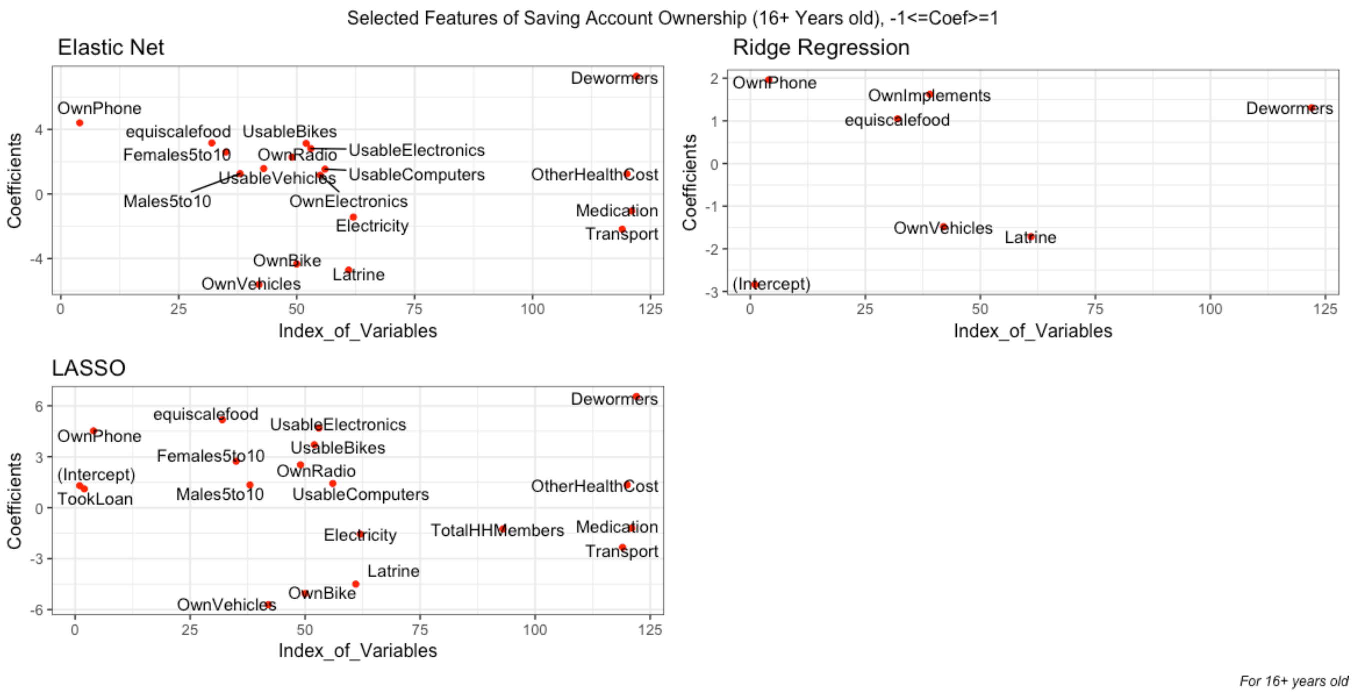
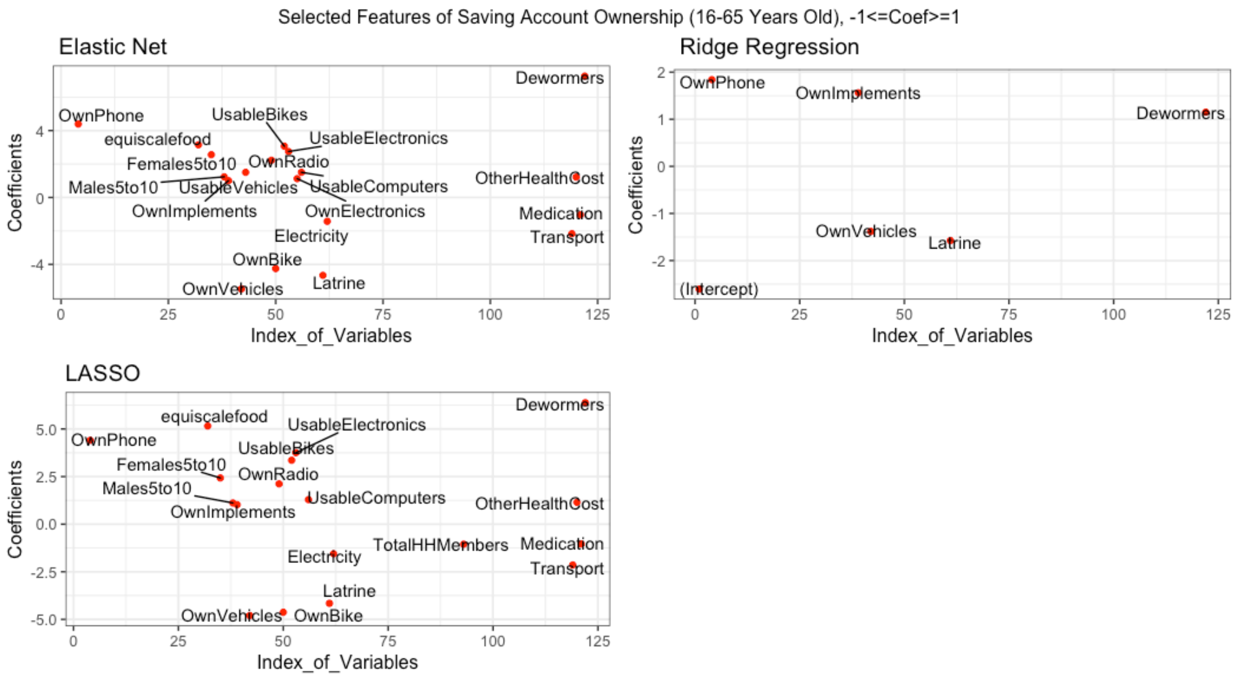
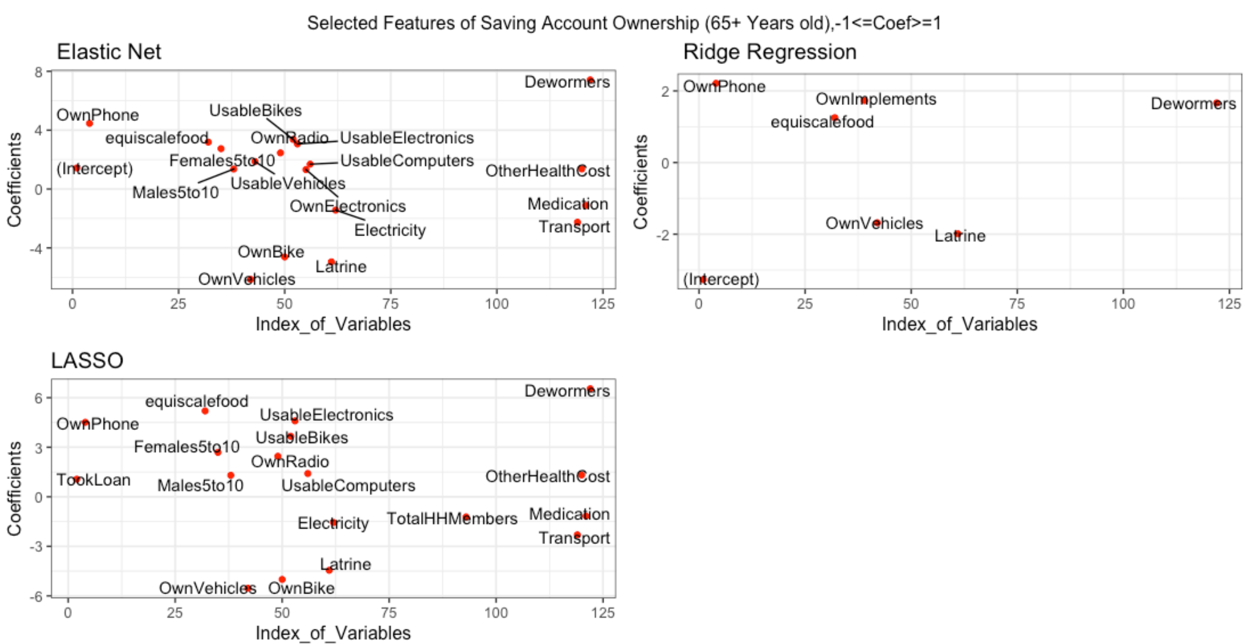


Figure 4: Top Predictors of Saving Account Ownership in Labor Force Age Group



For 16-65 years old

Figure 5: Top Predictors of Saving Account Ownership in Senior Age Group



For 65+ years old

Regardless of the regularization method, owning a mobile phone positively increases the probability of saving account ownership. In the Elastic net and LASSO models, keeping the other predictors equal to zero, the odds of saving account ownership are 11.74 times for people who own a mobile phone than for those who do not. In the ridge regression model, the odds of saving account ownership are six times higher for individuals who own a mobile phone than those who do not. The results are consistent with some slight deviations across age groups. The finding corroborates the finding in Mbiti and Weil (2016), Steinert et al., (2018), Nyoka (2018) Batista and Vicente (2020) who in different regions of Africa including in Kenya, have found that owning a mobile phone drives the take-up of mobile banking saving

The mobile banking has high penetration in Kenya, and that trend is growing as more people have access to smartphones. Supposedly, having a mobile banking account or being involved in mobile banking transactions (sending and receiving money or making payment) is a good proxy to own a saving account. We do not have the proportion of people owning a conventional banking saving account in comparison with those having a mobile banking saving account instead. Nevertheless, the adoption rate and market penetration of mobile banking products like Mpesa are more significant penetration than the conventional banking in Kenya. According to Hove and Dubus (2019) 74% of 15+ years old have access to mobile financial systems compared to 55.4% in traditional bank system or the same age category

Because mobile banking reduces the expropriation risk and increases the proprietary rights but also alter the bargaining power in and outside of the household (Jakiela and Ozier, 2012; Dupas and Robinson 2013b), it seems like an appropriate outlet for saving. In fact, owning a saving account on mobile banking reduces the peer and family pressure into spending and ultimately incentivizes the habit of saving of an individual. For, example, Dupas and Robinson (2013a) observe in their study that as the commitment device sophistication increases the saving account take up as well as the outcomes effects increase. That is to say, that an immaterial saving account that increases privacy would have at least the same if not a more significant effect than a lockbox, as used in Dupas' experimental study.

On the permanent income argument, our most regularized model does not feature the lagged quarterly income for any regularization parameter. This observation reinforces Deaton's argument against the use of the permanent income hypothesis in understanding household's saving behavior in developing countries. Instead, the uncertainty and the cyclicality of the income generation (on-farm cash income) are real features that might determine better household's saving account ownership.

4.2. Saving Accumulation Prediction

Figures 6-8 plot the most important features that have the highest marginal effect on the saving balance and the highest predictive weight.

Figure 6: Top Predictors of Saving Balance in Overall Sample

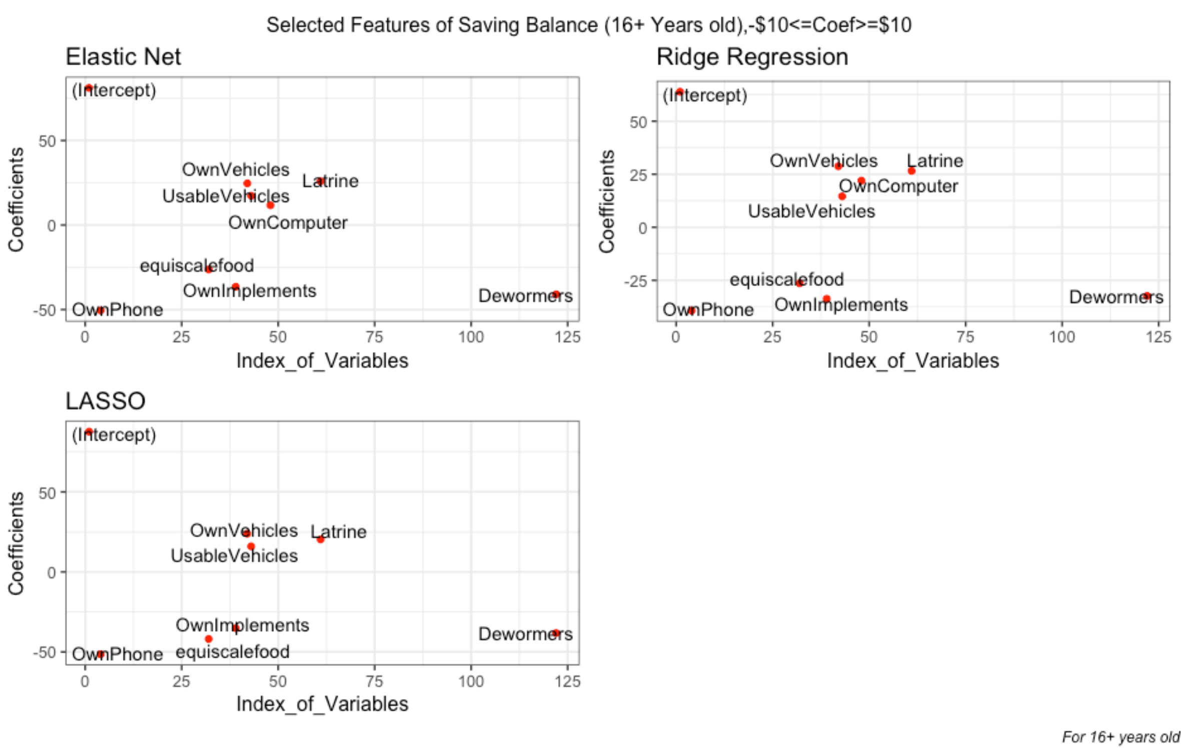


Figure 7: Top Predictors of Saving Balance in Labor Force Age Group

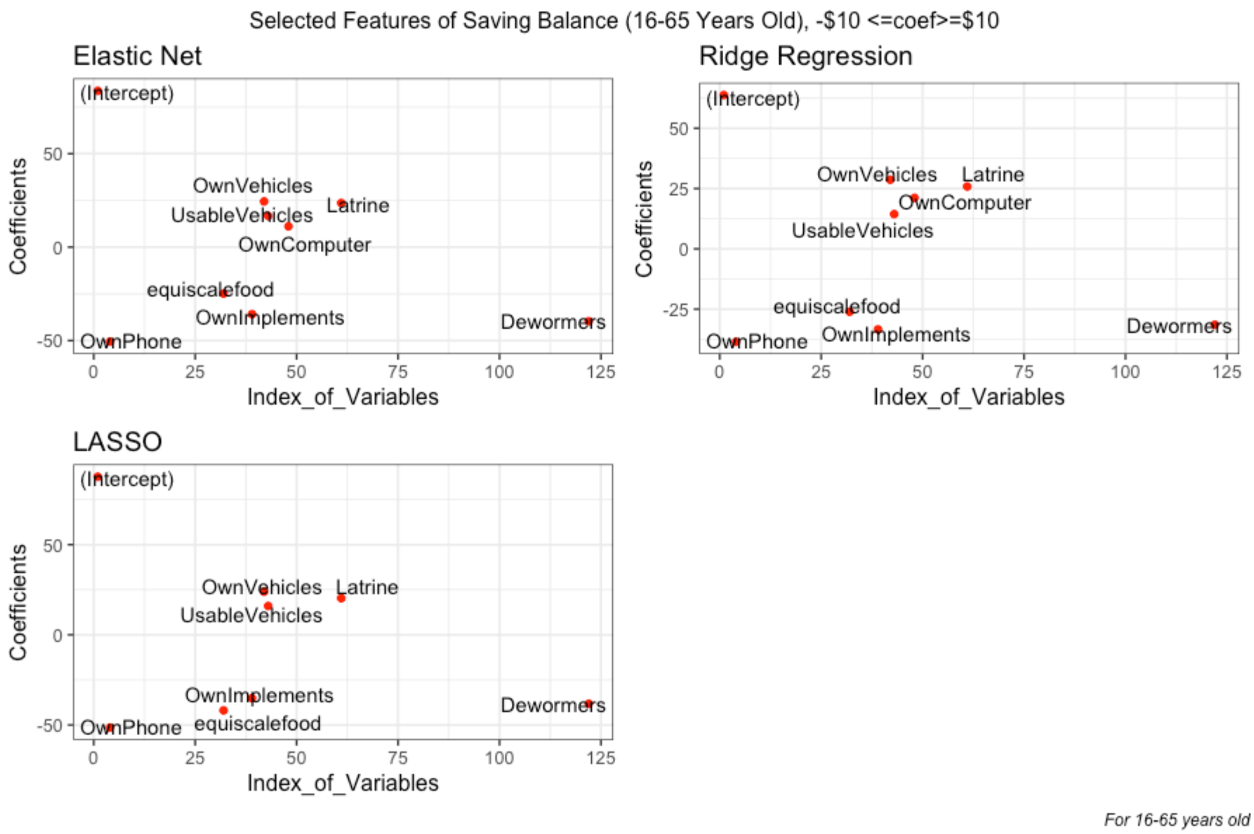
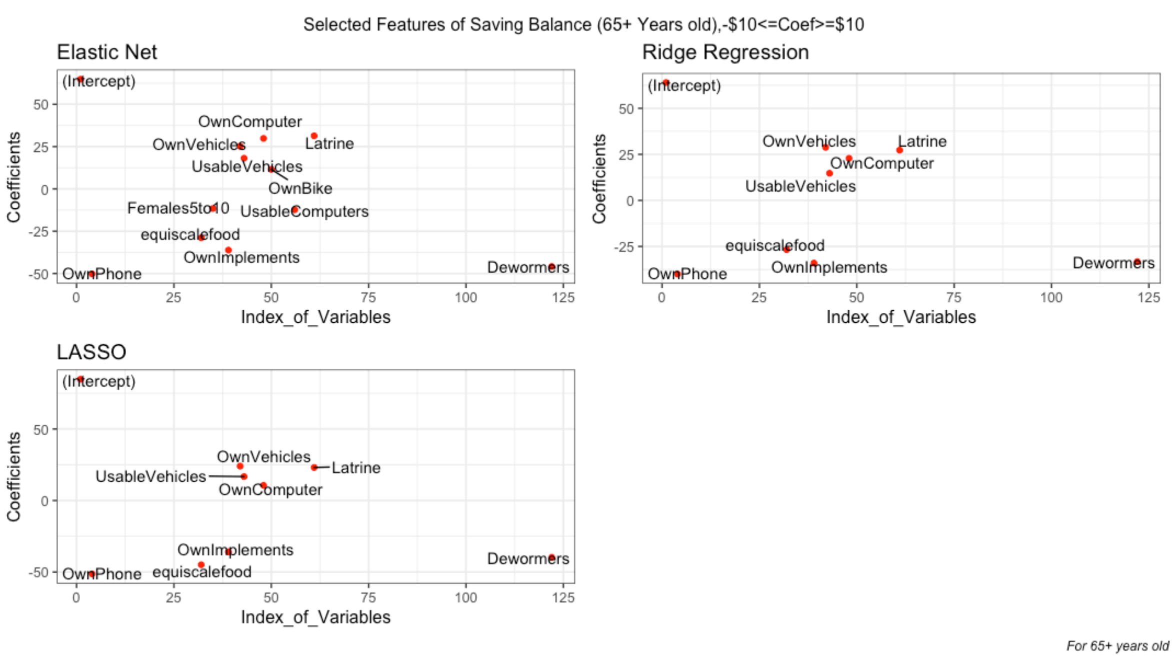


Figure 8: Top Predictors of Saving Balance in Senior Age Group



In average, the marginal effect of owning a mobile phone is about $50 less on the saving balance. This result is consistent across the regularized models we examine but also across the age groups. The results also confirm earlier findings (Mbiti and Weil, 2016; Ouma et al., 2017; Steinert et al., 2018; Nyoka, 2018; Batista and Vicente, 2020). For at least two reasons, this result might be intuitively explained. Because the mobile phone allows instantaneous access to the money, we could infer that if the saving account is a mobile banking account individual can withdraw and runs multiple transactions such as payments. Earlier papers (Ouma et al., 2017; Kibicho and Mungai, 2019; Seck Fall et al., 2020) have made a similar observation. Mobile banking is used in payments (electricity, for example) operations in Kenya. The second plausible explanation is related to the transaction costs of using a mobile banking account. Excepted the fees of using the service, there are no transactions costs due to distance, paperwork, or time opportunity cost (in waiting in line to withdraw money in conventional banks accounts). The transactional costs advantage -as explained in Dupas and Robinson (2013b), Prina (2013) and Schaner (2013) -may increase the number of transactions or even their amount. In that sense, for the sake of comparison, a lockbox will allow a fewer number of transactions if any and then a subsequent larger balance accumulation. Consider how less convenient it might be to open many times per week the lockbox to withdraw money for payments or investments, for instance. Finally, holding a saving account on mobile banking might have fewer commitment clauses such that the access to the money will be made conveniently and readily available than other forms of saving accounts.

A household head dissaves about $2 when she holds a saving account. However, if she belongs to the group of Elders, she tends to save about $2 more. When she has a salaried occupation, she saves $3 more. The characteristics of the household housing, the marital status of the adults in the household have limited marginal effects on the saving balance. Crop loss reduces the saving balance by $3.78, while the livestock loss has almost no marginal impact on the saving balance. The difference may stem from the weight of cropping in comparison with the livestock breeding in the income generation process of the households in that region. The living assets (livestock), as well as the tenureship of the land, have a limited contribution to the saving balance.

Depending on the regularization method, the size of the household marginally increases the saving balance in the rage of $0.5 to $ 3.13.The composition of the household, especially the number of female children between 5 and 10 years old, is associated with a dissaving of $7.18. However, the number of female children under age 5, marginally increases the saving balance by $2.39.

Across regularization methods and age, the permanent income marginal effects on the saving balance are limited. We do not find evidence of saving balance prediction role of lagged quarterly on-farm cash income. This result confirmed the finding in Deaton (1989) that rejects the permanent income hypothesis in many developing countries.

4.3. Age Groups Effects: A case for life cycle saving

That mobile phone ownership leads to higher mobile banking saving, and low accumulation in saving is consistent through age groups. However, one notable difference between labor force (16-65 years), seniors (65+ years) and the overall sample (16+ years) resides in the regularization methods outcomes and partially in the magnitude of the estimated coefficients. Compared to older individuals, mobile banking saving account ownership is attractive to a younger individual. The result is consistent with prior findings (Karakara and Osabuohein, 2019) who found that mobile banking saving account take-up is high in that age category relatively to older individuals. However, a young individual tends to dissave more compared to more aged individual once the former group adopts the saving product.

4.4. Performance of the Regularization Methods

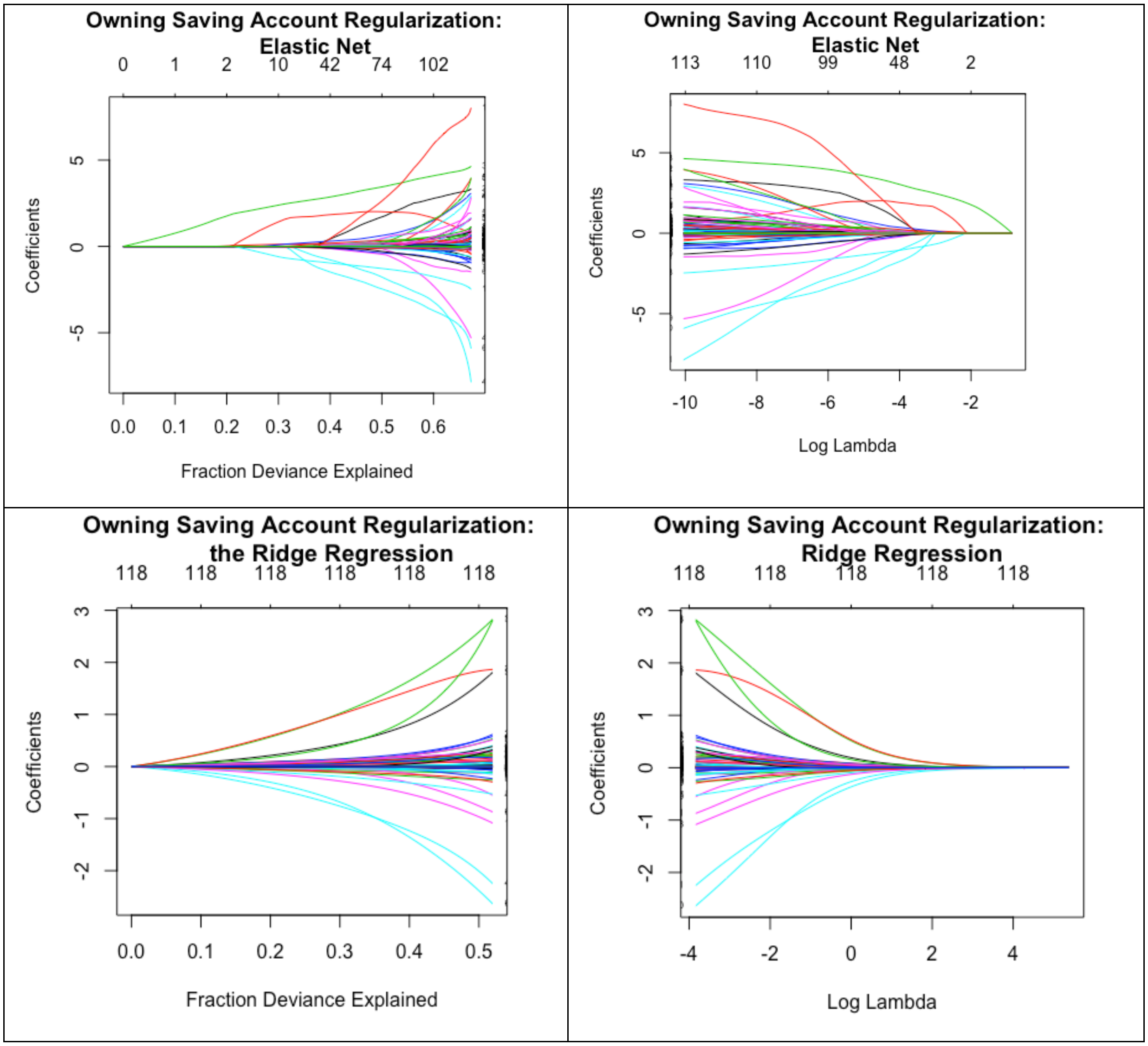
We report in table 2 the number of features selected for each regularized model.

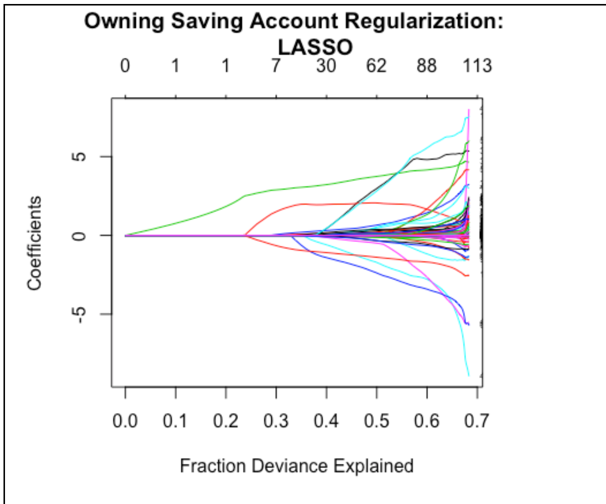
Elastic net and LASSO selects roughly a similar number of features even though the magnitude of the coefficients is slightly different for the selected features. In most cases, and across age groups, the two models select around 85-100 variables. However, as expected ridge regression selects a higher number of features than Elastic net and LASSO. For instance, up to 117 out of 123 variables are selected for the ridge regression in most cases. Moreover, the magnitude of the coefficients of the selected features is reasonably different. The grouping effect (by which an algorithm selects a group of highly correlated variables instead of just one in the group) exhibited by the elastic net is offset here for the simple reason that the number of observations is not greater than the number of predictors. For about 4,000 observations, the model has only 123 parameters. Figure 9 plots the fraction of the deviance explained and the log of lambda of each method for predicting both saving account ownership and balance.

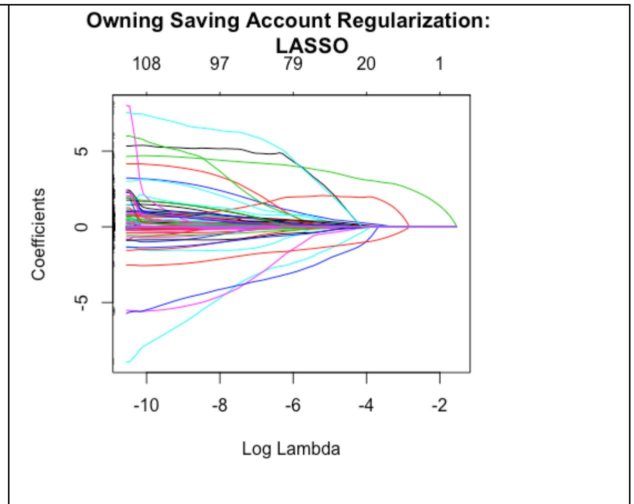
Table 2: Model Performance: Number of Selected Features for Saving Account Ownership and Accumulation Prediction

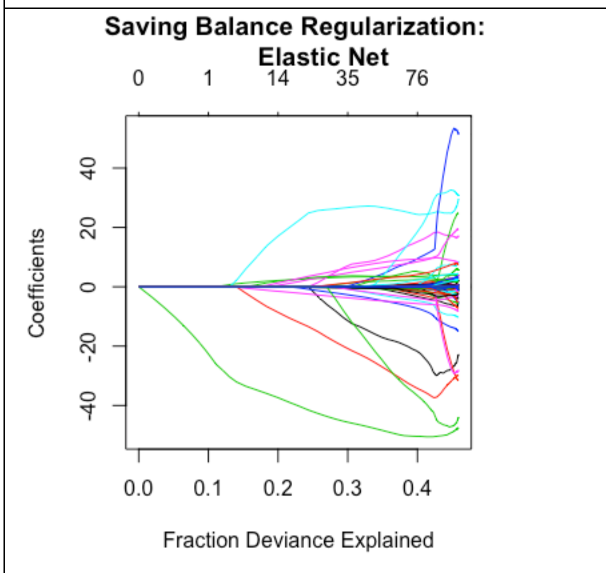
	16+ Years Old			16-65 Years Old			65+ Years Old		
	Elastic Net	Ridge Regression	LASSO	Elastic Net	Ridge Regression	LASSO	Elastic Net	Ridge Regression	LASSO
	Saving Account Ownership								
Number of Selected Features	107	117	98	111	117	99	106	117	95
	Saving Accumulation								
	73	117	94	81	81	78	101	117	75

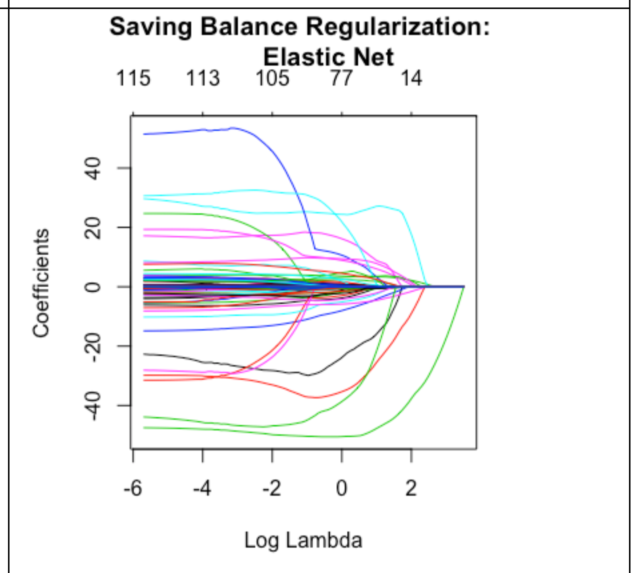
Figure 9: Deviance and Log Lambda of Each Regularized Regressions for Saving Account Ownership and Balance

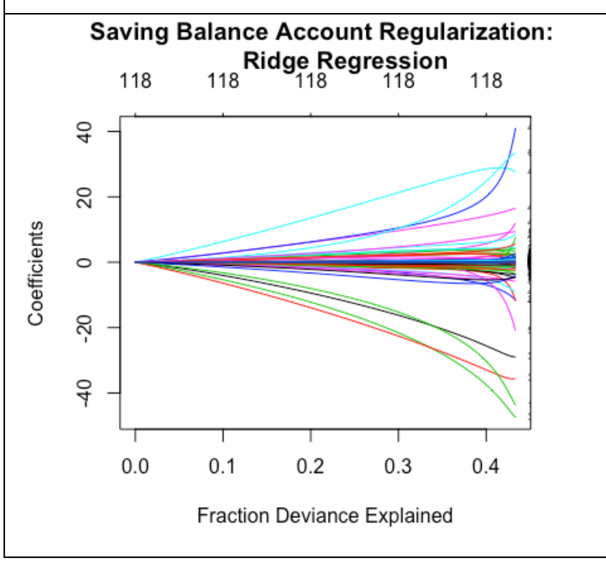


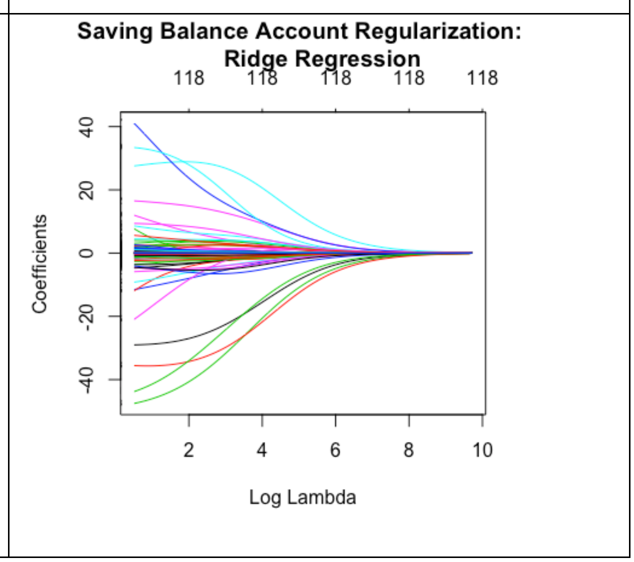

Owning Saving Account Regularization:
LASSO
0 1 1 7 30 62 88 113
Coefficients
5
0
-5
0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7
Fraction Deviance Explained

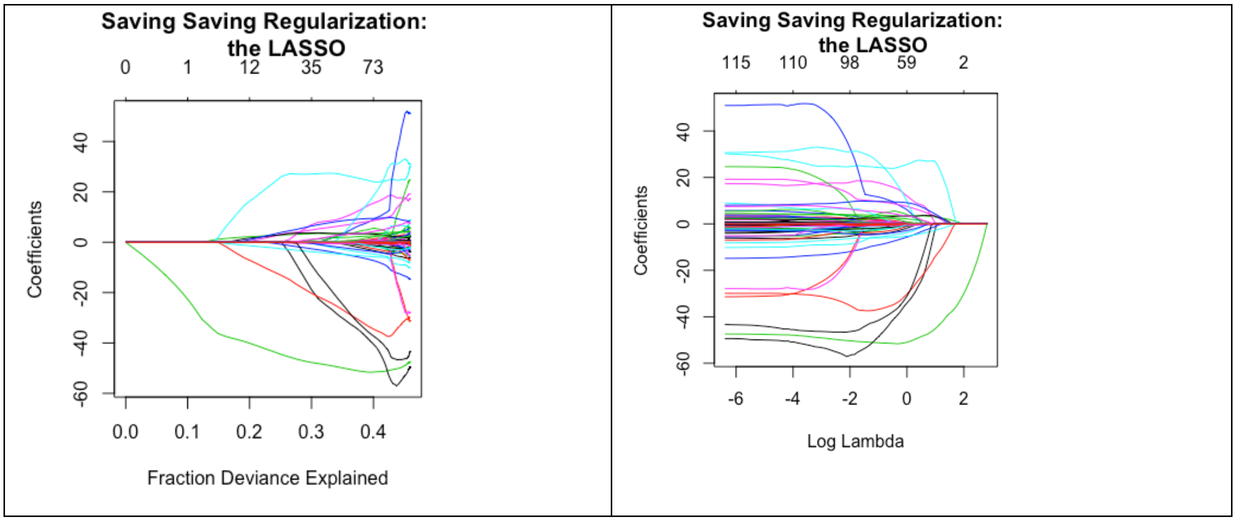

Owning Saving Account Regularization:
LASSO
108 97 79 20 1
Coefficients
5
0
-5
-10 -8 -6 -4 -2
Log Lambda


Saving Balance Regularization:
Elastic Net
0 1 14 35 76
Coefficients
40
20
0
-20
-40
0.0 0.1 0.2 0.3 0.4
Fraction Deviance Explained


Saving Balance Regularization:
Elastic Net
115 113 105 77 14
Coefficients
40
20
0
-20
-40
-6 -4 -2 0 2
Log Lambda


Saving Balance Account Regularization:
Ridge Regression
118 118 118 118 118
Coefficients
40
20
0
-20
-40
0.0 0.1 0.2 0.3 0.4
Fraction Deviance Explained


Saving Balance Account Regularization:
Ridge Regression
118 118 118 118 118
Coefficients
40
20
0
-20
-40
2 4 6 8 10
Log Lambda



5. Conclusion

Household's saving behavior in developing countries does not fit into the widely admitted theories to explain the saving pattern in developed economies. The permanent income hypothesis is challenged, and many behavioral economics findings might not necessarily provide a better explanation. While the surge in digitalization unlocked some promises of saving through mobile banking, financial inclusion leveraging these paths need a better characterization. At the same, algorithmic methods to study the subject is now possible with machine learning techniques, offering an opportunity to contrast and improve the outcomes of classical econometrics. That is the subject of this paper, that considers mobile banking saving take-up and saving accumulation through the lens of penalized regressions. We predict role of the devices and technology-based solutions in the saving account ownership and saving deposit amount using data from rural Kenya. We take a machine learning approach to choose the important features in predicting both saving account ownership and saving balance amount. The paper contributes to the methodological strides in studying household's saving behavior in developing countries, harnessing machine learning algorithms to reduce bias and overfitting in estimation.

We find that owning a mobile phone predicts saving account ownership, but individuals tend to dissave more when using the device. We confirm the limited role of the permanent income hypothesis, family structure and demographics in the saving pattern in general. The age group difference resides in the magnitude (the weight of predictors) of coefficients with the notable result that individual in the workforce tend to own a saving account but dissaves more. In contrast, if they happen to have a saving account, seniors tend to have higher saving balance. Our results shed light on the role of mobile banking in saving behavior in rural Africa. They convey information on plausible features that can guide the design of saving product to increase take-up of saving account and saving accumulation in rural Kenya.

References

Ashraf, N., G., N., Karlan, S.D., Yin, W. 2003. A Review of Commitment Savings Products in Developing Countries

Asongu S.A., J. C. Nwachukwu. 2016. African Finance Journal - Mobile Phone Penetration, Mobile Banking and Inclusive Development in Africa. *African Financial Journal*, 18 (1), Pages 34 – 52.

Batista C., and P. C. Vicente. 2020. Is Mobile Money Changing Rural Africa? Evidence from a Field Experiment.

Banerjee A, Duflo E. 2017. "The Economic Lives of the Poor" Journal of Economic Perspectives. 21:141–167.

Bewly, T., 1977. The permanent income hypothesis: A theoretical formulation. *Journal of Economic Theory*, 16 (2). Pages 252-292

Browning M., and A. Lusardi. 1996. Household Saving: Micro Theories and Micro Facts. *Journal of Economic Literature*, 34 (4). pp. 1797-1855

Brune L, Giné X, Goldberg J, Yang D. "Commitments to Save: A Field Experiment in Rural Malawi" 2013. University of Maryland Working Paper, available at: http://econweb.umd.edu/~goldberg/docs/bggy_mwisavings.pdf.

Caballero R.J. 1990. Consumption Puzzles and Precautionary Savings Journal of Monetary Economics 25(1) 113 – 136

Coleman, A. 1998. Household Savings: A Survey of Recent Microeconomic Theory and Evidence. Treasury working paper. New-Zealand

Deaton, A. 1989. "Savings in Developing Countries: Theory and Review." Proceedings of the World Bank, Annual Conference on Development Economics.' Washington, DC

Demirguc-Kunt A, Klapper L. "Measuring Financial Inclusion: The Global Findex Database" 2012. World Bank Policy Research Working Paper No. 6025, available at: http://www-wds.worldbank.org/servlet/WDSContentServer/WDSP/IB/2012/04/19/000158349_20120419083611/Rendered/PDF/WPS6025.pdf.

Dupas P, Green S, Keats A, Robinson J. 2012. "Challenges in Banking the Rural Poor: Evidence from Kenya's Western Province". NBER Working Paper No. 17851, available at: http://www.nber.org/papers/w17851.

Dupas, P. and Robinson, J. 2013a. Why Don't the Poor Save More? Evidence from Health Savings Experiments. American Economic Review 2013, 103(4): 1138–1171

Dupas P, Robinson J. "Savings Constraints and Microenterprise Development: Evidence from a Field Experiment in Kenya" American Economic Journal: Applied Economics. 2013b;5:163–19

Fernandes D, Lynch J, Netemeyer R. "Financial Literacy, Financial Education and Downstream Financial Behaviors" 2013. *Management Science*, forthcoming, available at: http://papers.ssrn.com/abstract=2333898.

Flavin, M. 1981. The Adjustment of Consumption to Changing Expectations about Future Income. Journal of Political Economy 89(5) 974- 1009.

Friedman, M. 1957. A Theory of the Consumption Function (Princeton: Princeton University Press)

Friedman, Jerome, Trevor Hastie, and Rob Tibshirani. 2010. "Regularization Paths for Generalized Linear Models via Coordinate Descent." Journal of Statistical Software, Articles, 33 (1): 1–22.

Hall, R.E and Mishkin, F. (1982) The Sensitivity of Consumption to Transitory Income: Estimates from Panel Data on Households. Econometrica 50(2) 461-481
Hall, R. E. 1978 Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence. Journal of Political Economy 86(6) 971-987

Hoerl, A. & Kennard, R. 1988. Ridge regression, in 'Encyclopedia of Statistical Sciences', Vol. 8, Wiley, New York, pp. 129–136.

Hove Leo Van and Antoine Dubus. 2019. M-PESA and Financial Inclusion in Kenya: Of Paying Comes Saving? *Sustainability 11* (568).

Jakiela P, Ozier O. "Does Africa Need a Rotten Kin Theorem? Experimental Evidence from Village Economies" 2012. World Bank Policy Research Working Paper, available at: http://pamjakiela.com/jakielaozier_20121106c.pdf

Karakara A.A., and E. Osabuohien. 2018. Households' ICT access and bank patronage in West Africa: Empirical insights from Burkina Faso and Ghana. *Technology in Society,* 56. Pages 116-125

Karlan D., A. L. Ratman, J. Zinman. 2014. Saving by and for the poor: research review and agenda. Review of Income and Wealth 60(1): 36–78.

Kibicho, N., & Mungai, J. (2019). Mobile Banking Adoption and Financial Credit Accessibility in Wote Sub – County, Makueni County, Kenya. *International Journal of Current Aspects*, *3*(IV), 65-79. https://doi.org/10.35942/ijcab.v3iIV.47

Ky, Serge Stéphane, Mobile Phone for Financial Inclusion: Mobile Money Adoption in the West African Economic and Monetary Union (May 24, 2019). Available at SSRN: https://ssrn.com/abstract=3481194 or http://dx.doi.org/10.2139/ssrn.3481194

Ky S., C. Rugemintwari and A. Sauviat. 2018. Does Mobile Money Affect Saving Behaviour? Evidence from a Developing Country Journal of African Economies, Volume 27, Issue 3, June 2018, Pages 285-320, https://doi.org/10.1093/jafeco/ejx028

Laibson, D. 1997. "Golden Eggs and Hyperbolic Discounting." The Quarterly Journal of Economics. May 1997

Lowenstein G. and R. H. Thaler. 1989. "Anomalies: Intertemporal Choice." Journal of Economic Perspectives. 3(4):181-193

Mbiti, I., and D. N. Weil. 2016. Mobile Banking: The Impact of M-Pesa in Kenya. Chapter in NBER book African Successes, Volume III: Modernization and Development, Sebastian Edwards, Simon Johnson, and David N. Weil, editors. University of Chicago Press (p. 247 - 293)

Modigliani, F. 1966. "The life-cycle hypothesis of saving, the demand for wealth, and the supply of capital." Social Research. 33, 160-217

Nyoka C. 2018. An Examination of the Factors that Determine Consumers' Adoption of Mobile Banking Services in South Africa. *Euro Economics*, 03, Pages 116-129.

O'Donahue, T, and Matthew R.. 1999. "Doing it Now or Doing it Later." The American Economic Review. 89(1): 103-121

Ouma, S. A., T.M. Odongo and M. Were. 2017. Mobile Financial Services and Financial Inclusion: Is it a Boon for Savings Mobilization? *Review of Development Finance* 7(1). Pages 29-35

Prina S. "Banking the Poor via Savings Accounts: Evidence from a Field Experiment" 2013. Case Western Reserve University, Weatherhead School of Management Working Paper, available at: http://faculty.weatherhead.case.edu/prina/pdfs/prina_savingsaccounts_2013.pdf.

Seck Fall F., L. Orozco and A-M Akim. 2020. Adoption and use of mobile banking by low-income individuals in Senegal. *Review of development economics.* https://doi.org/10.1111/rode.12658

Schaner S. “The Cost of Convenience? Transaction Costs, Bargaining Power, and Savings Account Use in Kenya” 2013c. Dartmouth College Working Paper, available at: http://www.dartmouth.edu/~sschaner/main_files/Schaner_ATM.pdf.

Simon, Noah, Jerome Friedman, Trevor Hastie, and Rob Tibshirani. 2011. “Regularization Paths for Cox’s Proportional Hazards Model via Coordinate Descent.” Journal of Statistical Software, Articles. 39 (5): 1–13

Simon, Noah, Jerome Friedman, and Trevor Hastie. 2013. “A Blockwise Descent Algorithm for Group-Penalized Multiresponse and Multinomial Regression.”

Skinner, J. 1988. Risky Income, Life Cycle Consumption, and Precautionary Savings. Journal of Monetary Economics 22 237 – 255

Steinert, J.I., J. Zenker, U. Filipiak, A. Movsisyan, L. D. Cluver and Y. Shenderovich. 2018. Do Saving Promotion Interventions Increase Household Savings, Consumption, and Investments in Sub-Saharan Africa? A Systematic Review and Meta-Analysis. *World Development*, 104, Pages 238-256

Thaler, R. H. 1990. “Anomalies: Saving, Fungibility and Mental Accounts.” Journal of Economic Perspectives. 4(1): 193-205

Thaler, R.H.1992. The Winner’ s Curse: Paradoxes and Anomalies of Economic Life. Princeton: Princeton University Press

Thumbi SM, Njenga MK, Marsh TL, Noh S, Otiang E, Munyua P, et al. (2015). Linking Human Health and Livestock Health: A “One-Health” Platform for Integrated Analysis of Human Health, Livestock Health, and Economic Welfare in Livestock Dependent Communities. PLoS ONE 10(3): e0120761. https://doi.org/10.1371/journal.pone.0120761

Tibshirani, R. (1996), ‘Regression shrinkage and selection via the lasso’, Journal of the Royal Statistical Society, B 58, 267–288.

Tibshirani, Robert, Jacob Bien, Jerome Friedman, Trevor Hastie, Noah Simon, Jonathan Taylor, and Ryan J. Tibshirani. 2012. “Strong Rules for Discarding Predictors in Lasso-Type Problems.” Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 74 (2): 245–66

Zeldes, Stephen P (1989) Optimal Consumption with Stochastic Income: Deviations from Certainty Equivalence. Quarterly Journal of Economics 275 – 298

Zins, A., and L. Weil. 2016. The determinants of financial inclusion in Africa. *Review of Development Finance*, 6 (1). Pages 46-57

Zou, H. and Hastie, T. (2005). "Regularization and variable selection via the elastic net." J. R. Statist. Soc. B, 67: 301–320.