Entrepreneurial Saving Practices and Reinvestment: Theory and Evidence *

Thorsten Beck †
Cass Business School, City University London
CEPR

Haki Pamuk ‡
Wageningen University

Burak R. Uras §
Tilburg University

November 17, 2016

Abstract

We use a novel enterprise survey to gauge the relationship between saving instruments and entrepreneurial reinvestment. We show that while most informal saving practices are not associated with a lower likelihood of reinvestment when compared with formal saving practices, there is a significantly lower association of *saving within the household* with the likelihood of reinvesting profits than other savings form, most importantly, formal saving forms. This result is robust to the model specification and controlling for a large array of variables including district-level fixed effects. We also provide empirical tests to address reverse causation and omitted variable concerns. Our work contributes to the recent debate on the implications of different saving instruments in developing countries and expands the entrepreneurial financing constraints literature by focusing on internal rather than external funding constraints.

Keywords: Entrepreneurial finance; savings; reinvestment; financial inclusion.

JEL Classification: D14; G21; O12; O16.

*We wish to express our gratitude to the Financial Sector Deepening Trust Tanzania (FSDT), for kindly providing us with the main data set used in the study. We would like to thank Daan van Soest, Erwin Bulte, Matreesh Ghatak, Michele Tertilt, Silvio Daidone, Benedikt Goderis, Gonzague Vannorenberghe, Bert Willems, Janneke Pieters, Christian Mpalanzi and FSD Tanzania, and seminar and conference participants at Tilburg Economics Workshop, 2014 Tilburg Development Economics Workshop, and CSAE Conference 2014 for valuable comments and suggestions. This research was funded with support from the Department for International Development (DFID) in the framework of the research project 'Co-ordinated Country Case Studies: Innovation and Growth, Raising Productivity in Developing Countries'. All remaining errors are ours.

†E-mail: TBeck@city.ac.uk

‡E-mail: haki.pamuk@wur.nl

§Corresponding Author. E-mail: r.b.uras@uvt.nl, Address: Tilburg University Department of Economics Room K414 5000 LE Tilburg, The Netherlands

1 Introduction

In developing countries, intermediation costs and enforcement frictions constrain entrepreneurs' access to external finance - leaving their retained earnings as a key element for small business growth. For instance, Beck, Demirguc-Kunt and Maksimovic (2008) document that small enterprises in developing countries finance on average at least 60% of their investment through earnings retention. Given the importance of internal financing for sustainable business growth in developing world, *saving for business purposes* is expected to be positively correlated with earnings reinvestment decisions. In this paper we suggest that the saving mechanism itself might be a critical element in determining the ability of entrepreneurs to reinvest. Given the limited access to formal saving services, in many developing countries, entrepreneurs use informal mechanisms of saving and liquidity management (Allen et al. (2005 and 2012); Collins et al. (2009)). Entrepreneurs saving via informal channels, however, are more likely to have limited access to their savings. For instance, members of ROSCAs cannot access their savings until their turn comes[1], unless there is a relevant secondary market (Calomiris and Rajamaran (1998)). Similarly, saving via household members might inhibit entrepreneurs from reinvesting, because of redistributive pressures and commitment problems.

In this paper, we utilize a novel entrepreneurial survey to explore whether entrepreneurial saving practices have an association with entrepreneurs' reinvestment decisions. Specifically, we gauge whether the decision to save with formal financial institutions, individually (under the mattress), within the household or via other informal arrangements, such as rotating savings and credit associations (ROSCAs), is related to the decision to reinvest entrepreneurial earnings. We motivate our empirical work with a simple theoretical model that shows that an entrepreneur's reinvestment decision depends on the entrepreneur's saving practice, in addition to productivity and borrowing capacity of her entrepreneurial firm. We show that entrepreneurs are more likely to invest in their businesses if they save in a fashion which allows them easy access to their funds, such as formal savings accounts or personal saving mechanisms.

To test the empirical relationship between saving patterns and entrepreneurial reinvestment decisions, we use a micro- and small enterprise (MSE) survey for over 6,000 entrepreneurs conducted in 2010 in Tanzania. The sample of entrepreneurs surveyed covers a large variety of enterprises in different locations and of different gender, educational profile and sectors. We document that entrepreneurs' saving *practices* do indeed co-vary with the likelihood of earnings retention at MSEs. The survey design allows us to differentiate between different savings vehicles, including within household saving, saving under the pillow, informal savings clubs, and formal deposit accounts. Our results show that the probability of reinvestment is significantly higher for savers and that when compared against formal deposit account holders, entrepreneurs with informal saving practices are significantly less likely to reinvest.[2] However, our analysis also shows that the type of informal saving instrument might be important in generating an adverse consequence for reinvestment. Specifically, we find that when we

[1] See Besley et al. (1993) for a theoretical discussion of ROSCAs).

[2] Our survey does not contain information on actual investment as we will discuss in more detail below.

compare the practice of keeping savings *within the household* against the practice of *having a deposit account at a formal financial institution*, the latter is more likely to be associated with reinvestment than the former. Other informal savers do not exhibit significantly different reinvestment rates when compared against formal savers.

While the cross-sectional nature of our data does not allow us to make complete claims on causality, we conduct a series of checks to ensure the robustness of our results to the inclusion of additional control variables, alternative model specifications, and endogeneity biases. First, stable coefficient estimates across an array of specifications show that it is unlikely that our results are driven by omitted variable bias even after including an array of additional variables - as proposed by Altonji et al. (2005). Second, in spite of the difficulty to identify truly exogenous determinants of saving patterns that affect reinvestment decisions not directly, we report instrumental variable regressions with the distance to the nearest bank as instrument.

While research on finance and investment has almost exclusively concentrated on the impact of external finance constraints on investment behavior, our paper adds to a very recent literature arguing that relaxing the internal finance constraints is vital for investment and enterprise growth in developing countries as much as allowing for access to external finance. In this respect, Banerjee and Duflo (2011) suggest that saving constraints in the developing world could imply poverty traps for low income households. Karlan, Ratan and Zinman (2013) state in their survey article that undersaving can have important welfare consequences, such as variable consumption, low resilience to shocks, and foregone profitable investments. Using experimental settings, Dupas and Robinson (2013) show positive effects of providing access to formal savings accounts on micro-entrepreneurs' investment in Kenya, while Brune et al. (2013) find in a study for Malawian cash crop farmers that using a commitment savings product increases investment and crop output.

Our paper makes several contributions to this new financial development research agenda, which focuses on *inefficiency of savings and real economic performance*. First, our paper shows theoretically and empirically that different informal saving practices among entrepreneurs in developing countries can have different implications for investment decisions. Second, we show that a large spectrum of informal saving mechanisms are not associated with a lower likelihood of entrepreneurial reinvestment relative to saving formally. Rather, we show a significant negative association of one particular informal saving method, namely, *saving within the household* with reinvestment likelihood, and highlight this inefficient saving practice as a channel through which formal financial development policies could induce entrepreneurial investment in developing countries.

Tanzania is a perfect setting to test the relationship between different saving practices and entrepreneurial investment decisions. Tanzania is a low-income country in East Africa, whose private sector is dominated by micro- and small enterprises. While the financial sector was liberalized in the 1990s and there is a large number of formal financial institutions, access to formal financial services is very low, with only 17% of adults having a formal bank account in 2011 (World Bank (2012)). Tanzania shares many characteristics with other low-income countries in Africa, including a very disperse

population and a high degree of informality.

This paper relates to several distinct literatures. First, our study relates to research on finance and entrepreneurial reinvestment, which has shown that entrepreneurs invest more of their profits if they expect higher private returns from their investment activity (e.g. Demirguc-Kunt and Maksimovic (1998)). Johnson et al. (2002) show the importance of property rights in explaining the earnings reinvestment among entrepreneurial firms across several Eastern European countries, while Cull and Xu (2005) show that both property right protection and access to external finance in the form of bank loans are associated with more reinvestment by Chinese entrepreneurs. We add to this literature by focusing on saving patterns as an additional factor to explain the variation in reinvestment decisions across micro- and small entrepreneurs.

Second, our paper relates to a growing literature on the relative importance of formal and informal financing mechanisms in developing countries (Allen et al. (2005) and (2012); Ayyagari et al. (2010); Chavis et al. (2011)) that has investigated the extensive use of informal financing sources, especially by small and young firms. We add to this literature by focusing on different informal savings vehicles. By this we contribute to understanding where and when the benefits might not exceed the large fixed costs associated with increasing the outreach of formal financial institutions (Beck, Demirguc-Kunt, and Martinez Peria (2007)).

Finally, our paper relates to the literature on barriers to *save* in developing countries (see Karlan, Ratan and Zinman (2013), for an overview). In addition to geographic, monetary and regulatory barriers, there are significant social constraints on saving behavior, partly related to the position of the entrepreneur within the household. Previous research has linked participation in informal savings clubs, such as ROSCAs, to intra-household bargaining problems (e.g., Besley et al. (1993), Anderson and Baland (2002)). We contribute to this literature by documenting how different informal savings vehicles are related to the likelihood of entrepreneurial reinvestment.

Before proceeding, we would like to address an important methodological issue. Unlike many other papers in this literature that discuss randomized control trials (RCTs), our paper relies on cross-sectional survey data and thus faces the usual endogeneity biases. While we address these concerns by using instrumental variables and by exploring the differential relationship between savings patterns and reinvestment decision across different entrepreneurial groups, we clearly recognize the limits of such cross-sectional analysis in drawing causal inference. However, our analysis allows a broader exploration of reinvestment decisions across a variety of different informal savings patterns and we complement our empirical analysis with a theoretical exploration of different savings mechanisms.

The rest of the paper is organized as follows. Section 2 presents a theoretical model to show how saving practices can influence entrepreneurial investment. Section 3 discusses the regression set-up. Section 4 presents the data. Section 5 discusses our main findings, while section 6 tests for reverse causality, and provides an extensive list of robustness checks. Section 7 concludes.

2 A 2-Period Model

In this section we develop a partial equilibrium heterogeneous firms model to study the interactions between entrepreneurial *business saving practices* and *profit reinvestment*. This allows us to formalize empirically testable hypotheses, which we confront with data in section 5.

2.1 Environment

There are two time periods, *1* and *2*; a continuum of entrepreneurs indexed by i; and a good - called *cash*. Cash can be invested, saved or consumed. Entrepreneurs have linear preferences:

$$U_i = c_{1,i} + \beta c_{2,i}, \tag{1}$$

where U is the life-time utility and c_1 and c_2 are consumption levels in period-1 and in period-2 respectively. The parameter β is a discount factor.

In the beginning of period-1, each entrepreneur is endowed with ω units of investable funds which we assume to be homogeneously distributed among all entrepreneurs in the economy. In period-1, entrepreneurs can utilize the ω-endowment to *re*-invest as productive capital ($k_{1,i}$), save as liquid reserves for business purposes ($s_{1,i}$), or consume ($c_{1,i}$), yielding the budget constraint for period-1:

$$k_{1,i} + c_{1,i} + s_{1,i} \leq \omega.$$

Each entrepreneur has access to a production technology that converts productive capital ($k_{1,i}$) into the output of period-2. The output realization of the entrepreneurial technology is conditional on a liquidity injection at the beginning of period-2.[3] Specifically, entrepreneur i's technology yields $A_i k_{1,i}$ units of cash in period-2 in return of $k_{1,i}$ units of capital investment in period-1 plus an additional $L(k_{1,i})$ if and only if the entrepreneur is capable of injecting $L(k_{1,i})$ units of funds at the beginning of the period-2 that is equal to or greater than $\ell_2 k_{1,i}$. In this production technology, the parameter $A_i > 1$ captures the productivity heterogeneity across entrepreneurs. The liquidity needs (L) can be financed via two sources:

(1): The entrepreneur can save cash from period-1 to period-2, which we will call *saving for business purposes* denoted by $s_{1,i}$, at a rate ζ_i with $\zeta_i \leq 1$. In this formulation, ζ_i captures the efficiency of the saving practice in transferring funds from one period to the other. We assume that there are two general saving practice types: Formal (bank deposits) (ζ_F) and informal (in various forms) ($\zeta_{I,i}$). We suppose that $\zeta_F = 1$ for those who save at a bank deposit account, whereas $\zeta_{I,i}$ is heterogeneous across agents. We assume, for instance, that "under the mattress" savers are more likely to keep the saved funds untouched by the remaining household members compared to those who save (and share)

[3] This can be thought of an additional working capital requirement for the effective use of the fixed asset investment, such as, e.g., utility costs.

together "with other household members".

(2): The entrepreneur can borrow, denoted with $b_{2,i}$, up to a θ_i fraction of $\ell_2 k_{1,i}$ in the financial market at a gross interest rate 1, where θ_i is an entrepreneur specific parameter capturing the ability to raise working capital finance externally. Formally:

$$\theta_i \ell_2 k_{1,i} \geq b_{2,i}.$$

Based on this specification, we can state the entrepreneurial total output at the end of the period-2 as follows:

$$\begin{aligned} y_{2,i} &= A_i k_{1,i} + L(k_{1,i}) \text{ if } L(k_{1,i}) \geq \ell_2 k_{1,i}, \\ &= L(k_{1,i}) \ if \ L(k_{1,i}) < \ell_2 k_{1,i}. \end{aligned} \tag{2}$$

The functional form (2) makes clear that if and only if the expected liquidity needs in period-2 can be financed, the capital available to entrepreneur i will yield the cash-flow generated by the technology, $A_i k_{1,i}$, in addition to the firm's liquidity holdings.[4] Hence, in this economy, firms must have the capacity to manage liquid reserves in order to be able to exploit productive investment opportunities. Entrepreneurs consume and repay using the period-2 output - yielding the following second-period budget constraint

$$c_{2,i} + b_{2,i} + L(k_{1,i}) \leq y_{2,i} + b_{2,i} + s_{1,i}\zeta_i. \tag{3}$$

Idiosyncratic parameters θ, A and ζ are drawn from a distribution function at the beginning of the period-1. Table 1 specifies the timing of events in both periods of the model.

- Table 1 about here -

2.2 Optimizing Behavior and Testable Hypotheses

Entrepreneurs maximize life-time preferences delineated at (1) subject to (2) and (3). An immediate implication of this model is that if and only if $k_{1,i} > 0$, the entrepreneur forecasts that there will be sufficient capacity to finance future liquidity needs. Therefore, the entrepreneur sets $k_{1,i} = 0$ if his capacity to finance liquidity is sufficiently low. This implies that as long as $k_{1,i} > 0$ we have two additional constraints that need to hold:

$$L(k_{1,i}) \leq s_{1,i}\zeta_i + b_{2,i}, \tag{4}$$

$$\theta_i L(k_{1,i}) \geq b_{2,i}. \tag{5}$$

[4]This type of a production function specification has been previously utilized in finance and development literature by Aghion et al. (2010): In their dynamic general equilibrium model, the authors introduce a complementarity between the ability to cope with future liquidity needs and current long-term investment and explain the negative correlation between volatility and growth observed in the cross-country data.

The qualitative properties of this model are then as follows. Entrepreneurs who choose a $k_{1,i} > 0$ exhaust their borrowing limit θ_i. This is implied by the assumption that carrying over cash is inefficient; or, more formally $\zeta_{I,i} < \zeta_F = 1 < A_i \ \ \forall i$. Hence, we have $b_{2,i} = \theta_i L(k_{1,i})$. Then using (4) with equality:

$$s_{1,i} = \left(\frac{1-\theta_i}{\zeta_i}\right) L(k_{1,i}). \tag{6}$$

Equation (6) implies that the lower ζ the higher is the amount of savings for business purposes - for those entrepreneurs who choose to invest. Using (6) in budget constraints (2) and (3) yields:

$$c_1 = \omega_i - k_{1,i} - \left(\frac{1-\theta_i}{\zeta_i}\right) L(k_{1,i}), \tag{7}$$

$$c_2 = A_i k_{1,i} + (1-\theta_i) L(k_{1,i}). \tag{8}$$

Letting the idiosyncratic rate of return from investment - or in other words the willingness to postpone consumption - be denoted with ρ_i, the optimal consumption plans implied by (1) are:

$$\begin{aligned} &c_{1,i} > 0, \ c_{2,i} = 0 \text{ if } \rho_i < \frac{1}{\beta}, \\ &c_{1,i} = 0, \ c_{2,i} > 0 \text{ if } \rho_i > \frac{1}{\beta}, \end{aligned} \tag{9}$$

where using (7) and (8) we can show that $\rho_i = \frac{A_i + (1-\theta_i)\ell_2}{1+\ell_2\left(\frac{1-\theta_i}{\zeta_i}\right)}$. The linear structure of the model implies that the entire *beginning of the period earnings* (ω) get utilized for the business via either capital re-investment ($k > 0$) or keeping liquid savings to satisfy working capital needs ($s > 0$) if and only if the efficiency of operating the entrepreneurial technology is large enough. Using (7) and (8), we solve for the unit production efficiency of capital for the entrepreneur, who chooses to re-invest earnings into her technology (ρ). We then compare the unit efficiency of operating the technology against the minimum required rate of return presented at equation (9); namely $1/\beta$, that would make the entrepreneur willing to choose $k > 0$ and $s > 0$ (and hence consume in period 2) instead of $k = 0$ and $s = 0$ (and consume in period 1).

In this model, capital re-investment ($k > 0$) refers to an *illiquid capital investment*, which cannot be used to satisfy *liquid working capital needs* in the second period. This is the reason why the entrepreneur saves liquidity to be utilized for business purposes ($s > 0$) and additionally demands external finance ($b > 0$) as long as $k > 0$. In this respect, the *total investment* in the model exceeds re-invested capital k (because of the presence of savings and external finance), but $s > 0$ and $b > 0$ are liquid funds and do not correspond to illiquid (fixed) capital investment of an entrepreneurial firm.

Finally, applying comparative statics at ρ_i, we capture the key empirically testable implications of the model as the following:

Hypothesis 1: Efficiency of saving practice (high ζ_i) increases the likelihood of reinvestment.[5]

Hypothesis 2: High borrowing capacity (high θ_i) increases the likelihood of reinvestment.[6]

Hypothesis 3: Productivity (high A_i) increases the likelihood of reinvestment.[7]

There is a close relationship between the variable k (capital reinvestment) and the variable s (saving for business purposes) in our model and specific survey questions that we will utilize to measure k and s. Specifically, we use the survey questions "Does 'I re-invest some of the profits back into the business?' apply to your business? (Yes/No)" to capture k and "Do you save for business purposes? (Yes/No)", to capture s.

3 The Empirical Methodology

We test the hypotheses derived from the theoretical model with a dataset collected from Tanzanian MSEs by the Financial Sector Deepening Trust of Tanzania. Specifically, we test whether entrepreneurs with more efficient saving patterns are more likely to reinvest. The reinvestment decision corresponds to the entrepreneurial decision in the theoretical model of investing the initial entrepreneurial endowment into physical capital. To test whether saving practices affect the decision to reinvest, we follow the literature and utilize a binary outcome variable *reinvest*, which equals one if the entrepreneur invests some of the profits back into business. We then estimate the following model

$$Prob\{Reinvest_i\} = \alpha + \gamma' S_i + \chi' Controls_i + \epsilon_i, \tag{10}$$

where i denotes the entrepreneur, S is a vector of saving practices comprised of dummy variable(s) which take(s) the value of 1 if the entrepreneur has the corresponding saving practice (see below for details) and ϵ is the error term. Since our dependent variable is binary, we estimate probit models for all different specifications of (10), and report marginal effects at mean levels for the coefficient estimates unless we state otherwise. The key empirical conjecture that we have in this analysis - based on hypothesis 1 that we derived in Section 2 - is that entrepreneurs with more efficient saving practices are more likely to reinvest.

We focus on reinvestment in our empirical analysis rather than investment for several reasons. First, while entrepreneurs are typically reluctant (or unable) to talk about specific amounts on investment or profits, it is easier to extract information on whether or not they have used their retained earnings for investment (Cull and Xu (2005)). For these reasons, our survey - as many other enterprise surveys - does not contain a reliable question on the actual amount of investment (to document quantities).

[5] Taking the first-partial derivative of ρ with respect to ζ we can see that $\frac{\partial \rho}{\partial \zeta} = \left(\frac{1}{\zeta^2}\right)\left\{\frac{\ell_2(1-\theta)(A+(1-\theta)\ell_2)}{\left[1+\ell_2\left(\frac{1-\theta}{\zeta}\right)\right]^2}\right\} > 0$.

[6] Defining $z \equiv \frac{1-\theta}{\zeta^2}$ and taking the first-partial derivative of ρ with respect to θ, we get $\frac{\partial \rho}{\partial \theta} = \frac{\frac{\ell_2}{z}(A-1)}{[1+\ell_2\zeta z]^2} > 0$.

[7] Taking the first-partial derivative of ρ with respect to A, we get $\frac{\partial \rho}{\partial A} = \frac{1}{1+\ell_2\zeta z} > 0$.

Therefore, earnings retention provides researchers often-times with a more reliable measure of investment. Second, our focus - both theoretically and empirically - concerns the capacity of entrepreneurs to self-finance. For our empirical work, we therefore prefer the reinvestment decision as dependent variable rather than investment (which might come from outside sources). This allows us to isolate the link between saving patterns and investment of entrepreneurs in their own company.

The vector of control variables included in the benchmark model is composed of an array of entrepreneurial and enterprise characteristics, partly informed by our theoretical analysis. First, we control for firms' past borrowing history. Specifically, *Borrowed* is a dummy variable which takes the value of one if the entrepreneur has ever borrowed to cover business needs, and is a proxy for the θ_i parameter in the theoretical model. Following testable hypothesis 2, businesses that have access to external finance are expected to reinvest more frequently.

Second, we use income level, education and business training history of entrepreneurs as proxies of entrepreneurial productivity A_i from our theoretical model. We conjecture that entrepreneurs with a higher *household income* can save more and as a result reinvest more often. To control for the income effects, we use self-reported monthly personal income levels.[8] Entrepreneurs with a high human capital are expected to be more committed to business growth, and to have higher rates of earnings retention. We therefore use the highest level of *formal education* completed by the respondents, as well as an indicator of entrepreneurial training, as this should matter for expected business performance and reinvestment behavior. According to hypothesis 3, proxies of productivity should be positively associated with reinvestment probability.

Third, although they are not discussed in our model, we additionally control for *gender* and *marital status* as previous studies showed that both can influence investment decisions (Iversen et al. (2006), Ashraf (2009), de Mel et al., (2009) and Fafchamps et al. (2013)). Specifically, we expect female entrepreneurs to face more claims on their income from spouse and family members. Similarly, married entrepreneurs might face more claimants on the business profits and might therefore be less likely to reinvest. Finally, we include sectoral dummies to control for sectoral performance that might explain reinvestment heterogeneity, as well as regional dummies to control for geographic heterogeneity in profitability, reinvestment and availability of different savings vehicles.

- Figure 1 about here -

We empirically explore the relationship between specific forms of savings and the likelihood of reinvestment. Figure 1 summarizes how we classify different saving practices. Specifically, our survey allows us to identify two types of saving practices among Tanzanian entrepreneurs:

1. *Save formal*: This practice includes the entrepreneurs who save their funds at formal financial institutions such as commercial banks, microfinance institutions or cooperatives.[9]

[8]Each respondent is asked which income range (e.g. TSHS 35 001 - TSHS 40 000 per month) describes their income level best. We use the median of that range (e.g. TSHS 37500.5) as the income level of the respondent.

[9]So entrepreneurs who save only formal and save both formal and informal (please see below for the definitions) are

2. *Save informal*: We consider entrepreneurs who do not save formally in this group.

This separation corresponds to the control-treatment group set-up of many randomized control trials that assess the impact of using formal saving products on household and entrepreneurial outcomes. In addition, however, our survey allows a finer classification to exploit the considerable heterogeneity in terms of informal saving practices. Therefore we first divide *save informal* into two groups:

1. *Save informal individually*: A large fraction of entrepreneurs in Tanzania save their funds only in a secret hiding place or a piggy bank, i.e. a cash container.

2. *Save informal with others*: This refers to practices of saving funds via informal saving clubs, such as ROSCAs, moneylenders or within household.[10]

To distinguish whether our entrepreneurs save through people living in the household or people who are not member of the same household, we decompose the practice of "*Save informal with others*" further:

1. *Save with household members*: The group comprises of entrepreneurs who give their funds to other household members to keep them safe.

2. *Save with people outside household*: The group contains entrepreneurs who save through ROSCAs or moneylenders.[11]

We again conjecture that entrepreneurs in the second group have more control over their savings than entrepreneurs in the first group, especially if the latter have limited intra-household bargaining power. In our regression analysis, we will use a dummy variable for each saving practice above (see Table 2 below for the descriptions) and work with different samples to compare both savers and non-savers but also different groups of savers in their reinvestment behaviour.

4 The Data

The survey data was collected by the Financial Sector Deepening Trust Tanzania in 2010 from a nationwide representative cross-section of 6,083 micro- and small enterprises.[12] The respondents of

considered in this group. Out of 597 formal savers, 186 also save informally, mostly individually. Only 19 formal savers also save with household members.

[10]Entrepreneurs who both save informal individually and save informally with others are considered in this group. Our results are robust when we create a separate dummy variable for this group having both saving practices and add them to the regressions.

[11]We do not include the entrepreneurs who both save informal with household members and save informal with people outside household in our analysis, because they are very few (only 7 observations) and classifying those in one group is difficult. Including those observations in either group do not change our results.

[12]Appendix B lists the relevant survey questions.

the questionnaire are entrepreneurs with an active business as of September 2010. Table 2 presents both detailed definitions of the variables and descriptive statistics of the sample.

- Table 2 about here -

The descriptive statistics in Panel A of Table 2 shows that the average number of employees among Tanzanian MSEs is 1.5 workers, ranging from one (i.e. self-employed) to 80 employees. [13] However, the median is one employee and 97% of entrepreneurs are self-employed. The average initial capital is about 35 USD and average monthly sales are 149 USD. The key question which we exploit to capture entrepreneurs' earnings retention asks whether *the respondent reinvests some of the profits back into business*. As we present in Table 2, 76% of the sample entrepreneurs engage in earnings retention. The sectoral breakdown in Panel B of Table 2 exhibits substantial variation: 54% and 30% of the businesses that operate in trade and service sectors, respectively, while 15% of enterprises that operate in manufacturing retain fractions of their earnings in their business.

Panel C of Table 2 presents characteristics of entrepreneurs and enterprises. About 50% of the entrepreneurs in the sample are female, 10% of the entrepreneurs are single. 30% of the sample entrepreneurs received business related training, and about 87% of the entrepreneurs have less than completed secondary education. 75% of the enterprises are located in rural areas. The median monthly personal income of entrepreneurs is 106 USD.[14]

Panel D of Table 2 presents descriptive statistics regarding the financing patterns of enterprises in our sample. Only 18% of all sample entrepreneurs ever borrowed for business purposes; 3% of entrepreneurs in the sample borrowed from a bank or MFI, 2% borrowed from a semi-formal financial institution, such as a SACCO or village bank and 6% borrowed from an informal source, such as money lenders, savings club or family and friends.

Saving is a common habit among the entrepreneurs in our sample. We utilize an extensive margin question asking whether the entrepreneur saves for business purposes, and distinguish savers from the rest of the population: 77% of the entrepreneurs in the sample save for business purposes. However there is considerable heterogeneity among *saving practices* of Tanzanian entrepreneurs. Informal individual saving is the most popular practice among Tanzanian entrepreneurs. 75% of the savers save informal-individually whereas only 13% of them save formally. Likewise, 13% of the savers do not save at a formal financial institution and instead save their funds via people outside the household such as members of ROSCAs and moneylenders or give them to household members.

Table 3 provides a uni-variate comparison between micro-entrepreneurs that reinvest and that do not reinvest. In line with our theoretical model, we find that borrowers and entrepreneurs with higher

[13]The relationship between business owners's saving and re-investment decisions might be weak in large businesses because of managerial layers. We test the robustness of our main result by excluding the businesses larger than 10 from our sample. Estimates reported in Table 2 do not change.

[14]This is computed with the average exchange rate for 2010. If using PPP exchange rates, the corresponding median income would be 288 dollars.

educational attainment and income are more likely to reinvest. Female entrepreneurs are less likely to reinvest. We do not find significant differences between entrepreneurs that do and do not reinvest in terms of marital status and training. Most importantly, entrepreneurs that reinvest are also more likely to save. This difference will be the focus of the following regression analysis, where we will dig deeper in terms of different savings vehicles, and their relationship with the likelihood of reinvestment.

- Table 3 about here -

Table OA1 in the Online Appendix presents a correlation matrix concerning the variables of interest for our analysis. The key variables such as "being a saver" and "retaining earnings within the business" exhibit a strong correlation. However, the sign of the relationship seems to be dependent on the saving practice of the respondents. In particular saving via others seems to be negatively correlated with firm reinvestment whereas formal and informal individual savers have higher reinvestment rates.

5 Saving Practices and Reinvestment: Main Results

Table 4 reports the marginal effects for the benchmark regression. We report heteroscedasticity robust standard errors in parentheses. Panel A reports regression results, while Panel B reports the difference in marginal effects between different groups of savers.

- Table 4 about here -

The results in the first column show that the probability of reinvestment is higher for both groups of savers compared to non-savers. Specifically, ceteris paribus, the reinvestment probability of an average Tanzanian MSE who saves informally is around six percentage points higher than for an entrepreneur who does not save, while the reinvestment probability of an average Tanzanian MSE who saves formally is around nine percentage points higher. The results in Panel B show a significant difference between formal and informal savers in the likelihood to reinvest. We also find that entrepreneurs with access to formal loans are more likely to reinvest, while formal business training increases the likelihood of reinvestment in business projects. Female and married entrepreneurs are less likely, while richer entrepreneurs are more likely to invest. Overall, these results are consistent with our theoretical predictions as discussed above and the existing literature.

In the next step, we study the implications of different saving *practices* on reinvestment. In order to test the predictions from our theoretical model, we rank saving *practices* based on their vulnerability to consumption temptations - as we discussed above - and investigate the implications of the variations in saving methods for the probability to reinvest. Specifically, we rank the "within household savers" as the group for whom the vulnerability to consuming savings is the highest. On the other extreme, we expect the most committed savers to be "formal savers" due to the highest opportunity cost of

consumption - resulting from the foregone interest income. Finally, comparing "informal individual savers" with "informal savers with others", we conjecture that while the redistributive pressure problem might be lower for the former, there would be a potential inflexibility to withdrawing savings when needed associated with the latter.

The results in column 2 of Table 4 show that entrepreneurs who save with others are seven percentage points less likely to reinvest than entrepreneurs who save formally, while entrepreneurs who save informally but individually are not significantly less likely to reinvest than "formal savers". Both formal savers and informal individual savers are more likely to reinvest than non-savers.

Finally, we focus on the group of respondents who save with others. We independently study the investment likelihood of household savers and respondents who save outside the household compared to the reinvestment probability of formal savers. The results in column (3) show that that *with household member savers* are ten percentage points less likely to reinvest than *formal savers*, a result significant at the 5% level. Furthermore, we also show that *savers with people outside household* are five percentage points less likely to reinvest than formal savers, though this result is significant only at the 10% level. There is no significant difference between *savers with household members* and *savers with people outside the household.* As before, formal and informal individual savers are more likely to reinvest than non-savers, while there is no significant difference between informal savers with others and non-savers.

In summary, our baseline empirical results are consistent with our theoretical model showing that inefficient saving practices are associated with a lower likelihood of reinvestment. They suggest that informal saving practices are associated with significantly lower likelihood of earnings retention compared to formal saving mechanisms. It is important to note that this finding is mainly driven by the difference in the reinvestment likelihood of within household savers and formal savers, for which the difference is most pronounced and statistically significant.

6 Robustness and Reverse Causality

While controlling for other enterprise and entrepreneurial characteristics reduces the risk that the relationship between savings patterns and the likelihood of reinvestment is a spurious one, we cannot exclude the possibility that our relationship is driven by different sources of endogeneity, including reverse causation and omitted variable bias. In this section, we present several robustness tests to mitigate concerns of reverse causation and omitted variable biases. We first conduct a series of checks to ensure the robustness of our key result to inclusion and exclusion of additional control variables and model specification before presenting instrumental variable estimates. We would like to stress that these different tests serve to mitigate concerns of endogeneity, but still do not allow us to draw clear causal inferences from our cross-sectional observational data.

6.1 Robustness tests

First, we address the potential mis-specification of our non-linear probit model. Table OA2 in the Online Appendix provides a set of regression results using Ordinary Least Square (OLS) estimation. Our coefficient estimates are consistent and significant, hence stable vis-a-vis our baseline probit regression estimates. Specifically, the OLS estimation shows that saving informally is negatively correlated with the likelihood of earnings retention compared to the formal saving mechanism, and this negative co-variance is significantly driven by the practice of saving with other household members.[15]

Second, we test the robustness of our key result to the inclusion of a vector of additional control variables, reported in Table OA3 in the Online Appendix. First, we add specific dummy variables for different sources of external finance at the start-up of the enterprise: formal, semi-formal and informal loans. However, none of the external financing variables have significant explanatory power for reinvestment likelihood. Second, we control for *entrepreneurial types* by utilizing the answers to the following survey question: "*why did you go to business?*"[16] As evidenced in the previous literature (Bruhn and Zia (2011)), transformational type entrepreneurs are expected to have higher rates of investment profitability and earnings retention rate compared to survival type entrepreneurs. Third, we add dummy variables to control for the type of the activity the business conducts. The activity of the business (e.g. buying and re-selling; buying, adding value and re-selling, providing a service etc.) may change the definition of reinvestment for business owner and timing of the reinvestment.[17] The estimates for the variables are jointly significant at the 1 percent level. Fourth, we include the logarithm of the initial start-up capital, the logarithm of current sales per employee, the logarithm of the duration of business and the logarithm of number of workers, since these size gauges are expected to determine the growth potential of a business- and hence the profitability of reinvestment. We also control for rural vs. urban location of the enterprise, as the accessibility to infrastructure might affect expectations and drive variations in reinvestment rates. Fifth, we control for number of businesses owned by the household by adding two dummy variables: 2 businesses and more than 2 businesses in the household. The survey respondents may save with other household members to reinvest to higher productivity businesses of other household members. So, the negative correlation between save within household and reinvestment may be due to multiple business ownership in the household.[18]

[15] In unreported robustness tests, we also use a matching model, focusing on the sub-group of formal savers and savers with household members. Doing this allows for a non-linear correlation between our explanatory variables and the decision to reinvest. Specifically, we match each entrepreneur who saves with household members with the three (nearest) counterfactual entrepreneurs who save formally - on the basis of the control variables listed in section 3 and column 3 of Table 4 - and confirm our previous findings.

[16] Entrepreneurs selected from a list of statements to indicate why they went into business. Multiple choices were available. The answers include: I was fired / lost/retrenched from a previous job; I couldn't find a job elsewhere; To support me / my family; To try out a business idea; I believe I can make more money working for myself than for someone else; I had nothing else to do/no other means of survival/no better option; parents / relatives were in business; I saw a good opportunity; I have always wanted my own business; I was encouraged by friends and relatives; I needed to supplement my income; Others, please specify.

[17] We include 5 separate dummy variables for the businesses buying and selling goods; buying, adding value and selling goods; making and selling goods; providing service; and other activities including agricultural ones.

[18] We also re-run the model for only the respondents who are members of households with one business. Our coefficient

Including all of these control variables does not affect our key empirical finding. We also replace the region fixed effects with district fixed effects to ensure that we are capturing geographical variations well enough that could explain the probability of reinvestment (column 2 of Table OA3). While our sample becomes smaller, our findings remain.[19]

Finally, as an additional check of omitted variable bias, we provide a stability test for our key regressions with saving with household members in the spirit of Altonji et al. (2005). Specifically, as Bellows and Miguel (2009) propose, we first estimate a parsimonious model and then add control variables gradually. The results in Table OA4 in the Online Appendix show that our marginal effect estimates in columns (1)-(6) are highly stable across different specifications. We measure the stability of marginal effects by calculating the ratio between the value in the regression including controls (numerator) - column (6) - and the difference between this effect and the one derived from a regression without covariates (denominator) - column (1). As Bellows and Miguel (2009) suggest, this ratio shows how strong the covariance between the unobserved factors explaining entrepreneurial reinvestment decision and saving practices needs to be, relative to the covariance between observable factors and saving practices, to explain away the entire effect we find. We compute this ratio for *save formally* and *save with people inside the household*. The ratio turns out to be -5, which suggests that to explain the full effect of save within household, the covariance between unobserved factors and saving practices needs to be more than five times as high as the covariance of the included control variables with saving practices.[20] We would like to highlight that in these regressions even adding region fixed effects does not significantly alter our coefficient estimates although adding regional dummies increases the pseudo R-square approximately four times, and regional dummies explain most of the variation in the model.

6.2 Instrumental Variable Analysis

In this subsection we investigate the relationship between saving choices and entrepreneurial characteristics, and then offer a test to alleviate endogeneity concerns. To investigate the determinants of saving choice, we replace the dependent variable *reinvest* with *save within household* in (10) and regress it on our list of control variables as well as one additional variable: distance to bank (section 2.4). The distance to the nearest bank is expected to reduce accessibility of "formal saving services". We use two distance to bank measures in our estimations. The first one is a subjective distance measure constructed by using the question from the enterprise survey: Is there any bank branch in one hour walking distance to your house? However, there might be a concern regarding the subjective measure, as entrepreneurs who search for formal saving instruments are also those who are more likely to know of the existence of a bank in the close proximity. Therefore, the correlation between the search

estimate for save with household members does not change and is statistically significant.

[19]Note that when we include district fixed effects the total number of observations in the regression decreases to 650 because some districts are excluded from the regression in Probit estimations due to perfect prediction. Our estimates are robust when we estimate the same model with OLS and do not lose any observations.

[20]In a similar set-up, Altonji et al. (2005) find a ratio of 3.55. The interpretation is that the larger the ratio, the lower the likelihood of unobservables to explain away the entire effect.

intensity and some unobserved characteristics may bias our results. For this reason, we also add an objective distance measure, the logarithm of ward level minimum distance to the closest bank branch, MFI or ATM in 2013 which we constructed using data from the Financial Services Map.[21]

Panel A of Table 5 reports the marginal effect from probit estimations for the saving practice choice including both subjective and objective distance measures. As we conjecture, the likelihood of saving with household members is higher when entrepreneurs are farther away from banks. The rest of the estimates are also in line with our theory. Entrepreneurs who have access to external finance and entrepreneurs with higher education, better training or high income are more likely to save formally. Finally, female entrepreneurs seem more likely to save in formal institutions - perhaps to escape from redistributive pressures.

- Table 5 about here -

To address the endogeneity concerns in the relationship between savings practices and entrepreneurial reinvestment decisions, we use an instrumental variable methodology which makes use of the determinants of saving practice choice. Since our dependent and main explanatory variables are binary, we use a system approach, and utilize entrepreneur's distance to the nearest bank measures as instruments in a nonlinear recursive bivariate probit model.[22] Specifically the model is formulated as follows:

$$Reinvest_i = \phi + \delta Savehousehold_i + \eta' Controls_i + \sigma_i, \tag{11}$$

$$Savehousehold_i = \lambda + \kappa' Z_i + \pi' Controls_i + u_i. \tag{12}$$

We assume that error terms σ_i and u_i are distributed via bivariate normal distribution. So, $E[\sigma_i] = 0$, $E[u_i] = 0$ and $cov[\sigma_i, u_i] = \mu$. We identify the system by using the vector Z which includes the distance to bank measures and use a similar set of controls as in the main specifications.[23]

There may be exogeneity concerns regarding the relationship between the instruments and reinvestment. Distance to bank measures might correlate with business opportunities and induce entrepreneurs to reinvest. In order to address these issues, in Panel B of Table 5 we test the exogeneity of our instruments. As the standard overidentification test for the 2SLS is not an option with Bivariate Probit estimation, we utilize an informal test procedure commonly applied in the empirical literature (e.g.

[21] We use data from the Financial Services Map for Tanzania. This data set gives geographic coordinates of bank branches, MFIs and ATMs in 2013 across Tanzania. We match these data with the existing geographic coordinates of the wards from which entrepreneurial data are collected. Then we calculate the distance of the wards to each financial unit and pick the minimum distance. The correlation between the subjective and objective measure of distance to the nearest bank branch is -33%.

[22] We also estimate the same model by using the 2SLS method. We have the same expected signs for the variables of interest but the coefficient estimates are bigger and imprecise as the variance increases. We believe this is because both the dependent and independent variables of interest are binary. Chibus et al. (2012) suggests 2SLS may give very different results and imprecise estimates if the number of observations is lower than 5000 (in our case it is 797).

[23] We do not use sector dummies in the bivariate probit estimations since our model does not converge. However, not using sector dummies does not change our results since our main results shown in Table 4 are robust when we do not control for them.

Egger et al. (2011); Booker et al. (2013)), where we introduce our instruments into the benchmark reinvestment regression and study the coefficient estimates. In Panel B of Table 5 we show that the coefficient estimates from Probit estimations of the instruments are not statistically significant when we include them as additional explanatory variables in our reinvestment regressions. This implies that the instruments are correlated with saving practices, but they do not co-vary with reinvestment. We also show that the estimated cross-correlation coefficient -which we denote with $\hat{\mu}$ - is not statistically significant in either estimation. Therefore, we cannot reject the null hypothesis that σ_i and u_i are uncorrelated or reinvestment is exogenous to saving practice choice, once we control for other entrepreneurial characteristics. We also test the joint significance of our exogenous variables in the first stage of our bivariate probit model: they are jointly significant at the 1 percent level (Chi-square=10 and p-value<0.001).

In Panel C of Table 5 we present the recursive bivariate-probit estimates by using the two different distance measures as our instruments. Also, Table OA5 in the Appendix shows detailed estimation results for the model, including the control variables. The coefficient estimate of *save with household member* remains negative although less significant than before (with a p-value=0.12). We also note that the estimates for the exogenous variables have the expected signs. The probability to save in the household decreases as the proximity to bank decreases.

In summary, our IV regression confirms our previous results. We realize the limits of drawing causal inference from cross-sectional data; however, we see this test as only one of several reported in this section to show that our findings are not explained by simultaneity bias.

7 Conclusion

Past research has identified several factors that are important for entrepreneurial investment in developing countries. In this study, we explored how different entrepreneurial saving practices - i.e. saving via formal financial institutions, individually (under the mattress), within the household or within informal arrangements, such as ROSCAs - are related with the likelihood of reinvestment. To this end, we used a novel survey data set collected from MSEs in Tanzania and distinguished multiple saving practices of entrepreneurs as well their earnings retention behaviour. We motivate our empirical research with a simple theoretical model that shows how different saving practices can influence investment decisions. Our theoretical and empirical results show that not only are formal savers more likely to reinvest - a finding already established in the literature, but that different informal savings patterns have different associations with the likelihood that micro-entrepreneurs reinvest their savings into their businesses. Specifically, entrepreneurs who save by giving funds to other household members are less likely to reinvest than formal savers. We address endogeneity concerns and also provide an extensive list of robustness checks that confirm our results.

Our findings suggest that the entrepreneurs who need to protect their savings from consumption

commitments of other household members may benefit most from the introduction of formal saving instruments in low income areas. Therefore, from a development policy perspective, targeting entrepreneurs who have low decision power in the household and facilitating their access to formal saving instruments could be thought as a priority. Our results have important implications for the interactions between enterprise performance and financial access as well. Enterprises that exploit reinvestment opportunities are expected to be more likely to sustain higher productivity levels and survive more often. Access to efficient saving mechanisms in this respect could be key to facilitate enterprise performance in financially developing societies.

References

Aghion, P., Angeletos, G.M., Banerjee, A., & Manova, K. (2010). Volatility and growth: Credit constraints and the composition of investment. *Journal of Monetary Economics*, 57 (3), 246-265.

Allen, F., Chakrabarti, R., De, S., Qian, J., Qian, M. (2012). Financing Firms in India. *Journal of Financial Intermediation*, 21, 409-445.

Allen, F., Qian, J., Qian, M. (2005). Law, Finance and Economic Growth in China. *Journal of Financial Economics*, 77(2005), 57-116.

Altonji, J., Elder, T., & Taber, C. (2005). Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. *Journal of Political Economy*.

Ashraf, N. (2009). Spousal control and intra-household decision making: An experimental study in the Philippines. *The American Economic Review*, 1245-1277.

Attanasio, O., Augsburg, B., De Haas, R., Fitzsimons, E., & Harmgart, H. (2011). Group lending or individual lending? Evidence from a randomised field experiment in Mongolia. *EBRD mimeo*.

Ayyagari, M., Demirguc-Kunt, A., Maksimovic, V. (2012). Firm innovation in emerging markets: the role of finance, governance, and competition. *Journal of Financial and Quantitative Analysis*, 46(06), 1545-1580.

Ayyagari, M., Demirguc -Kunt, A., Maksimovic, V. (2010). Formal versus Informal Finance: Evidence from China. *The Review of Financial Studies*, 23(8), 3048-3097.

Banerjee, A., and Duflo, E. (2011). Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty. *Public Affairs*.

Banerjee, A., Duflo, E., Glennerster, R., & Kinnan, C. (2013). The miracle of microfinance? Evidence from a randomized evaluation. *MIT mimeo*.

Beck, T., Demirg-Kunt, A., and Maksimovic, V. (2008). Financing patterns around the world: Are small firms different?. *Journal of Financial Economics*, 89(3), 467-487.

Beck, T., Demirguc-Kunt, A., & Martinez Peria, M. S. (2007). Reaching out: Access to and use of banking services across countries. *Journal of Financial Economics*, 85(1), 234-266.

Bellows, J. and E. Miguel (2009), War and Local Collective Action in Sierra Leone, *Journal of Public Economics* 93(11-12), 1144–57.

Besley, T., Coate, S., & Loury, G. (1993). The economics of rotating savings and credit associations. *The American Economic Review*, 792-810.

Booker, K., Sass, T. R., Gill, B., & Zimmer, R. (2011). The effects of charter high schools on educational attainment. *Journal of Labor Economics*, 29(2), 377-415.

Bruhn, M., & Zia, B. (2011). Stimulating managerial capital in emerging markets: the impact of business and financial literacy for young entrepreneurs. *World Bank mimeo*.

Brune, L., Giné, X., Goldberg, J., & Yang, D. (2011). Commitments to save: A field experiment in rural Malawi. *World Bank Policy Research Working Paper Series*.

Calomiris, C. W., & Rajaraman, I. (1998). The role of ROSCAs: lumpy durables or event insurance? *Journal of Development Economics*, 56(1), 207-216.

Collins, Darryl, Jonathan Morduch, Stuart Rutherford, and Orlanda Ruthven (2009) *Portfolios of the Poor: How the World's Poor Live on $2 a Day*, Princeton University Press.

Cull, R., Xu, L. C. (2005). Institutions, ownership, and finance: the determinants of profit reinvestment among Chinese firms. *Journal of Financial Economics*, 77(1), 117-146.

Chavis, L. W., Klapper, L. F., & Love, I. (2011). The impact of the business environment on young firm financing. *The World Bank Economic Review*, 25(3), 486-507.

Chiburis, R. C., Das, J., & Lokshin, M. (2012). A practical comparison of the bivariate probit and linear IV estimators. *Economics Letters*.

De Mel, S., McKenzie, D., & Woodruff, C. (2008). Returns to capital in microenterprises: evidence from a field experiment. *The Quarterly Journal of Economics*, 123(4), 1329-1372.

De Mel, S., McKenzie, D., & Woodruff, C. (2009). Are women more credit constrained? Experimental evidence on gender and microenterprise returns. *American Economic Journal: Applied Economics*, 1-32.

Demirguc-Kunt, A., & Maksimovic, V. (1998). Law, finance, and firm growth. *The Journal of Finance*, 53(6), 2107-2137.

Dupas, P, & Robinson, J., (2013a). Savings Constraints and Microenterprise Development: Evidence from a Field Experiment in Kenya. *American Economic Journal: Applied Economics* 5 (1): 163-192. doi:10.1257/app.5.1.163.

Dupas, P, & Robinson, J., (2013b). Why Don't the Poor Save More? Evidence from Health Savings Experiments. *The American Economic Review* 103 (4): 1138-1171.

Egger, P., Larch, M., Staub, K. E., & Winkelmann, R. (2011). The trade effects of endogenous preferential trade agreements. *American Economic Journal: Economic Policy*, 3(3), 113-143.

Fafchamps, M., McKenzie, D., Quinn, S. R., & Woodruff, C. (2011). When is capital enough to get female microenterprises growing? Evidence from a randomized experiment in Ghana (No. w17207). *National Bureau of Economic Research.*

Iversen, V., Jackson, C., Kebede, B., Munro, A., & Verschoor, A. (2011). Do spouses realise cooperative gains? Experimental evidence from rural Uganda. *World Development*, 39(4), 569-578.

Johnson, S., McMillan, J., Woodruff, C., (2002). Property Rights and Finance. *The American Economic Review*, 92 (5), 1335-1356(22).

Kaboski, J. P., & Townsend, R. M. (2005). Policies and Impact: An Analysis of Village Level Microfinance Institutions. *Journal of the European Economic Association*, 3(1), 1-50.

Kaboski, J. P., & Townsend, R. M. (2011). A Structural Evaluation of a Large-Scale Quasi-Experimental Microfinance Initiative. *Econometrica*, 79(5), 1357-1406.

Karlan, D., Ratan, A., & Zinman, J. (2013). Savings by and for the Poor: A Research Review and Agenda (No. 1027).

Karlan, D. & Zinman, J., (2010a.) Expanding Credit Access: Using Randomized Supply Decisions to Estimate the Impacts. *Review of Financial Studies*, 23 (1): 433-64.

Karlan, D. & Zinman, J., (2010b). Expanding Microenterprise Credit Access: Using Randomized Supply Decisions to Estimate the Impacts in Manila. *Unpublished*.

Pande, R., & Burgess, R. (2005). Do rural banks matter? Evidence from the Indian social banking experience. *American Economic Review*, 95(3), 780-795.

Robinson, J. (2012). Limited insurance within the household: evidence from a field experiment in Kenya. *American Economic Journal: Applied Economics*,4(4), 140-164.

Schaner, S. G. (2012). Do Opposites Detract? Intrahousehold Preference Heterogeneity and Inefficient Strategic Savings. *Unpublished manuscript, Dartmouth Coll., Hanover, NH*.

Vonderlack, R. M., & Schreiner, M. (2002). Women, microfinance, and savings: Lessons and proposals. *Development in Practice*, 12(5), 602-612.

Wahhaj, Z. (2010). Social norms and individual savings in the context of informal insurance. *Journal of Economic Behavior & Organization*, 76(3), 511-530.

Figure 1: Classification of saving practices

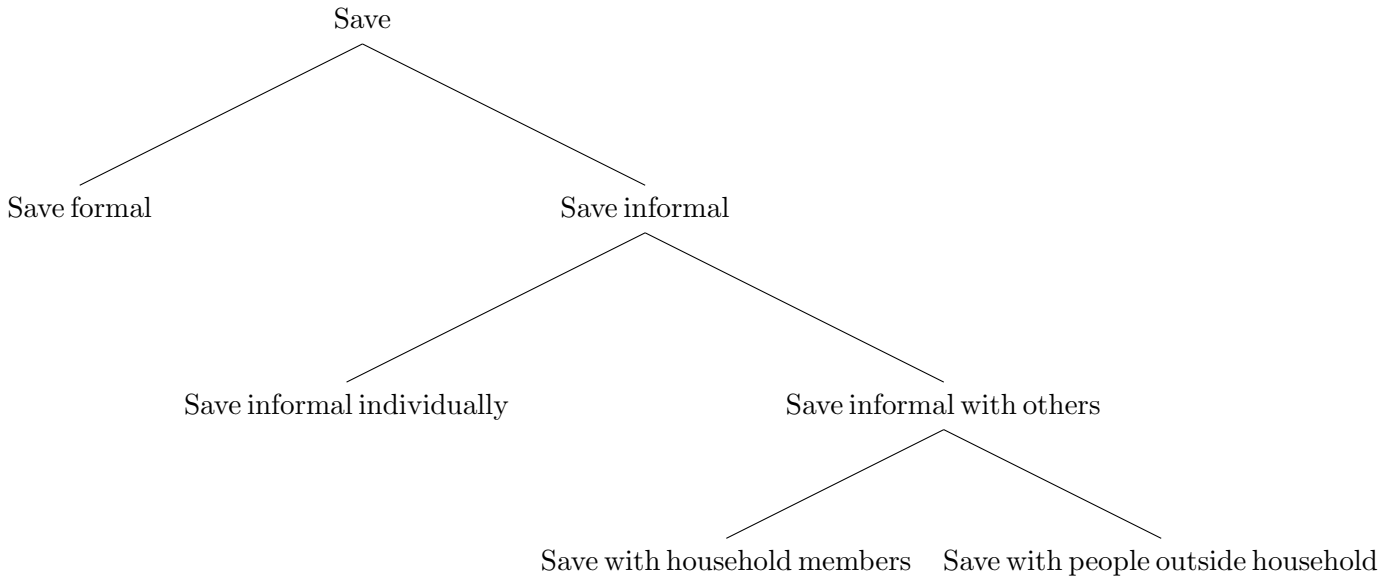

Table 1: Summary of model periods

I. Period-1

1. Entrepreneurial (3-dimensional) types are realized.
2. Capital investment into the production technology.
3. Saving for business purposes.
4. Period-1 consumption.

II. Period-2

1. Borrowing to finance liquidity needs.
2. Liquidity injection: Using borrowed funds and savings from period-1.
3. Cash-flow realization from the production technology.
4. Loan repayment.
5. Period-2 consumption

Table 2: Descriptive statistics for the main variables

Panel A: Firm characteristics	Description	Obs	Mean	S.D.	Min	Med.	Max
Reinvestment	Equals to 1 if respondent re-invest some of the profit back to business,	6083	0.76	0.43	0	1	1
Employee	Number of employees the business has (including owner)	6083	1.47	1.61	1	1	80
Initial capital	Logarithm of initial capital of the business, in Tanzanian Shillings	6083	10.62	2.21	0	10.82	25.33
Panel B: Sectoral breakdown of firms	Number of companies						
Trade	3291	54.1					
Service	1841	30.3					
Manufacturing	931	15.3					
Other	20	0.3					
Panel C: Entrepreneur characteristics	Description	Obs	Mean	S.D.	Min	Med.	Max
Education	Education level of the respondent, (0 none-6 university)	6077	2.00	0.89	0	2	6
Female	Equals to 1 if respondent is female,	6083	0.50	0.50	0		1
Single	Equals to 1 if respondent is single,	6083	0.10	0.29	0		1
No training	Equals to 1 if respondent has no business related training,	6083	0.70	0.46	0		1
Rural	Equals to 1 if respondent lives in a rural area,	6083	0.75	0.44	0		1
Income	Logarithm of personal income level of the respondent in Tanzanian Shillings	5868	11.94	1.15	9.90	11.92	15.20
Bank branch within one hour walking distance	Equals to 1 if there is a bank within one hour walking distance from the home of the respondent.	6083	0.30	0.46	0		1
Min. distance to ATM, bank branch, or MFI	Minimum distance of the ward, at which responts lives, to the nearest ATM, bank branch or MFI, (in logarithms and at ward level).	5580	2.01	1.79	-4.36	2.60	6.13
Age	Respondent age	6083	36.84	10.58	16	35	91
Panel D: Finance variables	Description	Obs	Mean	S.D.	Min	Med.	Max
Save	Equals to 1 if respondent saves for business purposes.	6083	0.77	0.42	0		1
Save formal	Equals to 1 if the respondent saves in a bank account, MFI or SACCO.	6083	0.10	0.30	0		1
Save informal	Equals to 1 if respondent saves but not in a bank account, MFI or SACCO.	6083	0.68	0.47	0		1
Save informal individually	Equals to 1 if the respondent saves in a secret hiding place or piggy bank and does not save formally.	6083	0.57	0.49	0		1
Save informal with others	Equals to 1 if the respondent saves with household members or save with people outside household and does not save formally.	6083	0.10	0.30	0		1
Save with household members	Equals to 1 if the respondent saves via by giving it to a household member to keep it safe and does not save formally.	6083	0.06	0.23	0		1
Save with people outside household	Equals to 1 if the respondent saves via by giving it to a non household member or merry go-round and does not save formally.	6083	0.05	0.21	0		1
Borrowed	Equals to 1 if the respondent has ever taken a loan/ borrowed money for business purposes.	6083	0.18	0.38	0		1
Formal loan	Equals to 1 if the respondent took a loan from a bank or MFI to set up or take over the business	6083	0.03	0.16	0		1
Semi formal loan	Equals to 1 if respondent took a loan from an employer, SACCO Village Bank, local government schemes or donor/NGO to set up or take over the business	6083	0.02	0.13	0		1
Informal loan	Equals to 1 if the respondent took a from family, friends, savings club, money lender or supplier to set up or take over the business	6083	0.06	0.24	0		1

Table 3: Comparison of company characteristics for reinvesting and not reinvesting businesses

	Reinvestment=1	Reinvestment=0	Difference
Borrowed	0.19	0.14	0.05***
	(0.39)	(0.34)	
Education	2.04	1.88	0.15***
	(0.88)	(0.92)	
Female	0.48	0.56	-0.07***
	(0.50)	(0.50)	
Single	0.10	(0.09)	0.01
	(0.30)	(0.28)	
No training	0.70	0.72	-0.02
	(0.46)	(0.45)	
Income	12.03	11.68	0.35***
	(1.16)	(1.09)	
Save	0.80	0.71	0.09***
	(0.40)	(0.45)	

Notes: This table summarizes the difference between the characteristics of the respondents that reinvest or do not reinvest. Column 1 and Column 2 reports listed variables'sample averages for reinvesting and not reinvesting respondents respectively. Column 3 reports the difference between the sample means. We report standard errors in parenthesis.* p<0.1. ** p<0.05. *** p<0.01

Table 4: Estimates for reinvestment and saving/saving practices relationship

	(1) reinvest	(2) reinvest	(3) reinvest
Panel A: Marginal effect estimates			
Save formal	0.09***	0.09***	0.09***
	(0.02)	(0.02)	(0.02)
Save informal	0.06***		
	(0.01)		
Save informal individually		0.07***	0.07***
		(0.01)	(0.01)
Save informal with others		0.01	
		(0.02)	
Save with household members			-0.01
			(0.02)
Save with people outside household			0.04
			(0.03)
Borrowed	0.04**	0.04**	0.04**
	(0.02)	(0.02)	(0.02)
Education	0.01	0.01	0.01
	(0.01)	(0.01)	(0.01)
Female	-0.03**	-0.03**	-0.03**
	(0.01)	(0.01)	(0.01)
Single	0.04**	0.04**	0.04**
	(0.02)	(0.02)	(0.02)
No training	-0.03***	-0.03***	-0.03***
	(0.01)	(0.01)	(0.01)
Income	0.03***	0.03***	0.03***
	(0.01)	(0.01)	(0.01)
Observations	5,803	5,803	5,803
Panel B: Comparison of formal and informal saving			
Save informal − Save formal	-0.03*		
	[0.07]		
Save informal individually − Save formal		-0.02	-0.02
		[0.18]	[0.18]
Save informal with other − Save formal		-0.07***	
		[0.00]	
Save with household members − Save formal			-0.10***
			[0.00]
Save with people outside household − Save formal			-0.05*
			[0.07]

Notes: This table shows the baseline estimation results for the relationship between saving practices, control variables and reinvestment likelihood. The detailed variable definitions are given in Table 2. Reinvestment is the dependent variable in the estimations. In Panel A, we report report marginal effects estimates at mean values from Probit estimations and robust standard errors in parentheses. To control for unobserved regional and sector level fixed effects, we add sector and and region dummies to all estimations. In Panel B, we report the differences between coefficient estimates for formal and informal saving practices, which are shown in Panel A. We conduct a Wald test to test whether the differences equal to zero and report p-values from that test in brackets. * p<0.1. ** p<0.05. *** p<0.01

Table 5: Instrumental variable estimation

Panel A: Estimates for save informally with household members vs. save formal	
Dependent variable:	Save informally with household members
Bank branch within one hour walking distance	-0.10**
	(0.05)
Min. distance to ATM, bank branch, or MFI	0.03*
	(0.01)
Borrowed	-0.35***
	(0.04)
Education	-0.12***
	(0.02)
Female	-0.09**
	(0.04)
Single	-0.02
	(0.07)
No training	-0.00
	(0.04)
Income	-0.11***
	(0.02)
Panel B: Exogeneity check	
Dependent variable:	Reinvestment
Save with household members	-0.10***
	(0.04)
Bank branch within one hour walking distance	-0.01
	(0.03)
Min. distance to ATM, bank branch, or MFI	-0.01
	(0.01)
Panel C: Bivariate probit estimates	
Dependent variable:	Reinvestment
Save with household members	-0.19$^{\perp}$
	(-0.12)
$\hat{\mu}$	0.22
	(0.29)
Observations	797

Notes: This tables reports estimates related to instrumental variables estimation.The sample used for estimations include only formal savers and informal savers with household members. We additionally control for region fixed effects by adding region dummies in all estimations. The number of observations at columns 1 and 2 are lower due to missing coordinate information for some wards. We control for Borrowed, Education, Female, Single, No training, Income, and region dummies in all models. The details for each Panel are as follows. Panel A shows the first stage estimates for the determinants of the Save informal with household members v.s. Save formal choice. Save informal with household members is the dependent variable, and Save formal is the base category in the estimations. We report marginal effect estimates at mean values from Probit estimations and robust standard errors are in parentheses. Panel B informally tests the exogeneity of our instruments: Bank branch within one hour walking distance and Minimum distance to ATM, bank branch, or MFI. If the estimates are not statistically significant than it implies that they do no have direct impact on reinvestment but have an impact o reinvestment only through saving practice choice. Reinvestment is the dependent variable in the estimation. We report marginal effect estimates at mean values from Probit estimations and robust standard errors in the parentheses. In Panel C, we report the coefficient estimate for save informal with household members from Bivariate Probit estimation. Robust standard errors clustered at ward level are in parenthesis. Our excluded instruments for saving practice choice between Save with household members and Save formal are Bank branch within one hour walking distance and Minimum distance to ATM, bank branch, or MFI. We report marginal effect estimates at mean values and clustered robust standard errors at ward level in the parentheses. μ is the correlation estimate for the error terms from model 15 and 16. $^{\perp}$ p<0.15, * p<0.1, ** p<0.05, *** p<0.01.

Online Appendix A: Tables

Table OA1: Correlation Matrix: Pairwise correlations among selected variables

	Reinvestment	Save	Save formal	Save informal individually	Save informal with others	Save with household members	Save with people outside household	Borrowed	Education	Female	Single	No training	Income
Reinvestment	1												
Save	0.09**	1											
Save formal	0.07**	0.18**	1										
Save informal individually	0.05**	0.62**	-0.38**	1									
Save informal with others	-0.03*	0.18*	-0.11*	-0.39*	1								
Save with household members	-0.04**	0.14**	-0.08**	-0.29**	0.74**	1							
Save with people outside household	0.00	0.12**	-0.07**	-0.25**	0.64**	-0.03**	1						
Borrowed	0.06**	0.12**	0.30**	-0.09**	0.03**	-0.01	0.05**	1					
Education	0.07**	0.13**	0.24**	-0.04**	0.02	0.02	0.01	0.17**	1				
Female	-0.06**	0.00	-0.05**	0.03**	0.00	-0.10**	0.11**	0.03**	-0.10**	1			
Single	0.02	0.00	0.00	0.01	-0.01	0.00	-0.01	-0.04**	0.07**	0.01	1		
No training	-0.02	-0.06**	-0.08**	0.02	-0.04**	-0.04**	-0.01	-0.10**	-0.12**	0.00	0.00	1	
Income	0.13**	0.12**	0.19**	-0.02	0.01	0.02	-0.02	0.12**	0.19**	-0.24**	-0.02	0.00	1

Notes: The table shows the pairwise correlation coefficient among selected variables. The detailed variable definitions are given in Table 2. ** indicates the estimate is statistically significant at least 5 percent level.

Table OA2: OLS estimates for reinvestment and saving/saving practices relationship

	(1) reinvest	(2) reinvest	(3) reinvest
Panel A: Marginal effect estimates			
Save formal	0.10***	0.09***	0.10***
	(0.02)	(0.02)	(0.02)
Save informal	0.06***		
	(0.01)		
Save informal individually		0.07***	0.07***
		(0.01)	(0.01)
Save informal with others		0.02	
		(0.02)	
Save with household members			-0.01
			(0.03)
Save with people outside household			0.04
			(0.03)
Borrowed	0.04**	0.04***	0.04**
	(0.01)	(0.01)	(0.01)
Education	0.01	0.01	0.01
	(0.01)	(0.01)	(0.01)
Female	-0.03**	-0.03**	-0.03***
	(0.01)	(0.01)	(0.01)
Single	0.04**	0.04**	0.04**
	(0.02)	(0.02)	(0.02)
No training	-0.03***	-0.04***	-0.04***
	(0.01)	(0.01)	(0.01)
Income	0.03***	0.03***	0.03***
	(0.01)	(0.01)	(0.01)
Observations	5,803	5,803	5,803
R^2	0.14	0.14	0.14
Panel B: Comparison of formal and informal saving			
Save informal − Save formal	-0.03*		
	[0.07]		
Save informal individually − Save formal		-0.02	-0.03
		[0.22]	[0.21]
Save informal with other − Save formal		-0.08***	
		[0.00]	
Save with household members − Save formal			-0.10***
			[0.00]
Save with people outside household − Save formal			-0.05*
			[0.08]

Notes: This table test the robustness of baseline estimation results reported in Table 4 by using OLS estimation. The detailed variable definitions are given in Table 2. Reinvestment is the dependent variable in the estimations. In Panel A, we report coefficient estimates from OLS estimations and robust standard errors in parentheses. To control for unobserved regional and sector level fixed effects, we add sector and and region dummies to all estimations. In Panel B, we report the differences between coefficient estimates for formal and informal saving practices, which are shown in Panel A. We conduct a Wald test for whether the differences equal to zero and report p-values from that test in brackets. * p<0.1. ** p<0.05. *** p<0.01

Table OA3: Estimates for reinvestment and saving practices relationship - additional control variables

	(1) reinvest	(2) reinvest
Panel A: Marginal effect estimates		
Save formal	0.08***	0.09***
	(0.02)	(0.02)
Save informal individually	0.06***	0.08***
	(0.01)	(0.01)
Save with household members	-0.00	-0.01
	(0.02)	(0.03)
Save with people outside household	0.03	0.04
	(0.03)	(0.03)
Formal Loan	0.04	
	(0.04)	
Semi formal loan	-0.10*	
	(0.06)	
Informal loan	0.02	
	(0.03)	
Initial capital stock	0.01*	
	(0.00)	
Sales per worker	0.02***	
	(0.01)	
Rural	0.01	
	(0.02)	
Size	0.04***	
	(0.02)	
Duration	0.01	
	(0.01)	
Having 2 businesses	0.06***	
	(0.02)	
Having more than 2 businesses	0.11***	
	(0.04)	
Observations	5,753	5,527
Entrepreneurial Dummy	Yes	No
Activity Dummy	Yes	No
Region FE	Yes	No
District FE	No	Yes
Panel B: Comparison of formal and informal saving		
Save informal individually − Save formal	-0.02	-0.01
	[0.23]	[0.23]
Save with household members − Save formal	-0.08***	-0.08***
	[0.00]	[0.00]
Save with people outside household − Save formal	-0.05	-0.04*
	[0.11]	[0.08]

Notes: In this table we test the sensitivity of our estimates presented in Column 3 of Table 4 by controlling for additional variables. The detailed variable definitions are given in Table 2. Reinvestment is the dependent variable in the estimations. We report marginal effects estimates at mean values from Probit estimations and robust standard errors in parentheses. Reinvestment is the dependent variable in all estimations. In Panel B, we report the differences between coefficient estimates for formal and informal saving practices, which are shown in Panel A. We conduct a Wald test to test whether the differences equal to zero and report p-values from that test in brackets. In all estimations we control for productivity proxies, education, training, and income but do not report estimates to economize on space. p¡0.01 * p<0.1. ** p<0.05. *** p<0.01

Table OA4: Test for the stability of estimates

	(1) reinvest	(2) reinvest	(3) reinvest	(4) reinvest	(5) reinvest	(6) reinvest
Panel A: Marginal effect estimates						
Save formal	0.13***	0.12***	0.09***	0.09***	0.09***	0.09***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Save informal individually	0.08***	0.08***	0.07***	0.07***	0.07***	0.07***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Save with household members	0.01	0.01	-0.01	-0.02	-0.02	-0.01
	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.02)
Save with people outside household	0.06**	0.05**	0.05*	0.06**	0.05**	0.04
	(0.02)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)
Borrowed		0.04**	0.02	0.03*	0.03	0.04**
		(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Education			0.02**	0.01**	0.01*	0.01
			(0.01)	(0.01)	(0.01)	(0.01)
Female				-0.04***	-0.04***	-0.03**
				(0.01)	(0.01)	(0.01)
Single				0.02	0.02	0.04**
				(0.02)	(0.02)	(0.02)
No training			-0.00	-0.00	-0.00	-0.03***
			(0.01)	(0.01)	(0.01)	(0.01)
Income			0.04***	0.04***	0.04***	0.03***
			(0.01)	(0.01)	(0.01)	(0.01)
Observations	5,803	5,803	5,803	5,803	5,803	5,803
Pseudo R-Squared	0.01	0.01	0.02	0.03	0.03	0.13
Sector FE	No	No	No	No	Yes	Yes
Region FE	No	No	No	No	No	Yes
Panel B: Comparison of formal and informal saving						
Save informal individually − Save formal	-0.05***	-0.04***	-0.02	-0.02	-0.02	-0.02
	[0.01]	[0.01]	[0.23]	[0.23]	[0.22]	[0.18]
Save with household members − Save formal	-0.12***	-0.12***	-0.11***	-0.11***	-0.10***	-0.10***
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Save with people outside household − Save formal	-0.07***	-0.05***	-0.05	-0.05	-0.05	-0.05*
	[0.01]	[0.01]	[0.17]	[0.26]	[0.24]	[0.07]

Notes: In this table, we test the stability of our estimates for saving variables by inserting key control variables separately. Reinvestment is the dependent variable in the estimations. We estimate Probit models for all specifications. In Panel A we report marginal effects estimates at mean values for all variables and robust standard errors in parentheses. In Panel B, we report the differences between coefficient estimates for formal and informal saving practices, which are shown in Panel A. We conduct a Wald test for whether the differences equal to zero and report p-values from that test in brackets. Robust standard errors in parentheses. The definitions of the variables are given in Table 2. The details of the estimations in the columns 1 to 6 are as follows: (1) In column 1, we estimate the model without a control variable. (2) In column 2, we control for access to finance by using our borrowing proxy (Borrowed). (3) In column 3, we control for productivity proxies, education, training, and income in addition to borrowing. (4) In column 4, we additionally insert Female and Single variable to control for gender and being married. (5) In column 5, we control for sectoral heterogeneity by using sector dummies . (6) In column 6, we finally insert region dummies to control for regional heterogeneity. *** p<0.01 * p<0.1. ** p<0.05. *** p<0.01

Table OA5: Bivariate Probit Estimates

Dependendent variable:	(1) Reinvestment	(2) Save with household members
Save with household members	-0.19$^{\perp}$	
	(0.12)	
Bank branch within		-0.10**
one hour walking distance		(0.05)
Min. distance to		0.03*
Atm, bank branch, or MFI		(0.01)
Borrowed	0.02	-0.39***
	(0.05)	(0.13)
Education	-0.01	-0.12***
	(0.02)	(0.02)
Female	-0.07**	-0.09*
	(0.03)	(0.05)
Single	0.06	-0.02
	(0.05)	(0.07)
No training	-0.08***	-0.00
	(0.32)	(0.04)
Income	0.01	-0.11***
	(0.02)	(0.02)
Constant	0.40	5.84***
	(1.34)	(0.84)
Observations	797	797

Notes: This table shows the detailed bivariate probit estimates for Panel C of Table 5. We report marginal effect estimates at mean levels for all estimations. Robust standard errors clustered at ward level are in parentheses. We use the sample for Formal Savers and Household Savers in all estimations. We additionally control for region dummies in the estimations. The details of columns 1 and 2 are as follows: (1) In column 1, we present the bivariate probit estimates of model (15) which is jointly estimated with model (16) using distance instruments. (2) In column 2, we present the bivariate probit estimates of model (16) including distance measures and jointly estimated with model (15). $^{\perp}$ p<0.15, * p<0.1, ** p<0.05, *** p<0.01.

Appendix B: Survey questions

1. Does "I re-invest some of the profits back into the business?" applies to your business? (Yes/No)
2. Including yourself, how many full-time employees did the business employ when it started operations? _____ (Record number)
3. What was the initial capital of this business? _____ (TSh).
4. What type of business was it? 1. Manufacturing 2. Retail 3. Services 4. Wholesale 5.Agriculture 6.Agricultural processing 7. Other specify
5. Highest level of formal education completed. (1 = none; 2= Preprimary 3= some primary; 4= primary completed; 5 = some secondary; 6= secondary completed; 7 = technical training after secondary ; 8= university)
6. Gender? (Male/Female)
7. Did you have any relevant training before the start of this business? (1. No training 2. Yes, business training 3. Yes, technical training)
8. Please give me the letter that best describes the average TOTAL MONTHLY PERSONAL INCOME before tax and other deductions. Please include all sources of income i.e. salaries, pensions, income from investment, etc. (A. Below TSHS 40 000 per month; B. TSHS 35 001 TSHS 40 000 per month; C. TSHS 40 001 TSHS 60, 000 per month; D. TSHS 60 001 TSHS 80, 000 per month; E. TSHS 80 001 TSHS 100 000 per month; F. TSHS 100 001 TSHS 200 000 per month; G. TSHS 200 001 TSHS 300 000 per month; H. TSHS 300 001 TSHS 400 000 per month; I. TSHS 400 001 TSHS 500 000 per month; J. TSHS 500 001 TSHS 1,000 000 per month; K. TSHS 1,000, 001 -1,500,000 per month; L. TSHS 1,500, 001 -2,000,000 per month; M. TSHS 2,000, 001 -2,500,000 per month; N. TSHS 3,000, 001 -3,500,000 per month; O. Over 3,5000,000 per month; P. Refuse to answer; Q. Uncertain/Dont know)
9. (filled in by interviewer) How old is the respondent? _____ (Record in years)
10. Where do you save? (1. In a bank account 2. Saving with a SACCO 3. Saving in a MFI 4. Give to a household member to keep safe 5. Give to somebody else to keep safe (non-household member) 6. Keep it in a secret hiding place 7. Put it in merry go round 8. Others specify)
11. As a small business owner, have you ever taken a loan/ borrowed money for business purpose? (1. Yes 2. No)
12. Where did you get the loan to set up or take over the business? Multiple Responses Possible (1. Loan from bank. 2. Loan from micro finance institution (MFI) 3. Loan from an employer 4. Loan from a SACCO 5. VICOBA (Village bank) 6. Loans from local Government /government schemes 7. Loan from friends and/or family 8. Loan from savings club 9. Loan from money lender 10. Loan/credit from the supplier 11. Loan from donor / NGO. 12. Other)