

Microcredit Loan Repayment Default among Small Scale Enterprises: A Double Hurdle Approach

Benjamin Musah Abu

University for Development Studies, Wa Campus, Ghana
abmusah@uds.edu.gh

Paul Bata Domanban

University for Development Studies, Wa Campus, Ghana
and

Haruna Issahaku

University for Development Studies, Wa Campus, Ghana
DOI//<http://dx.doi.org/10.4314/gjds.v14i1.8>

Abstract

Previous studies that analysed loan defaults have imposed the restrictive assumption that the factors which influence the probability of loan default and the rate of default are the same and with the same direction effects. This study introduces a two-stage approach by applying the double hurdle model to analyse probability of loan default and the magnitude of the default as a way of relaxing this restrictive assumption. Using the model, this study is able to isolate the determinants of default rate from the determinants of the probability of default. The control function approach is used to test and control for potential endogeneity of loan amount and profit in the probability of default equation. The model is applied to primary data from a survey of 200 small scale entrepreneurs in the Upper West Region of Ghana. Results show that enterprise size, interest rate, loan duration, level of profit and loan amount are the simultaneous determinants of probability and rate of default. The study recommends that the National Board for Small Scale Industries should intensify training programmes aimed at enhancing the managerial and technical capabilities of entrepreneurs to ensure healthy and productive management to promote enterprise growth and increase profit margins and hence reduce defaults.

Keywords: Repayment, Default, Microfinance, Small Scale Enterprises, Endogeneity, Double Hurdle

Introduction

Credit risk management is crucial to the survival of banks, microfinance institutions and related financial institutions, and the clients they serve. Credit risk can be analysed from three main angles: (1) the likelihood of default, (2) the loss given default –defined as one minus the recovery rate in the event of default, and (3) the degree of exposure in the event of default. The bulk of the empirical literature has focused on modelling the probability of default while little attention has been paid to modelling recovery rate and its relationship with default (Altman, Resti & Sironi, 2004). Recently, there is a new wave of research on recovery rates and its link with the probability of default. This new research effort notwithstanding, there are still some unresolved issues. First, it is still unclear whether the probability of default and the recovery rate are interdependent or independent (see McQuown, 1997; Wilson, 1998). The first set of literature which took root in the 1990s is based on the assumption that default probability and recovery rate are independent. Studies based on this assumption model recovery rate either as a constant term or as a stochastic variable separate from the probability of default. The stochasticity of the recovery rate implies that its volatility constitutes idiosyncratic risk which can be eliminated through proper portfolio diversification.

In contrast, the second set of literature holds that the recovery rate is related to the probability of default and that the correlation between the two is negative (see Frye, 2000; Jarrow, 2001; Altman, Brady, Resti & Sironi, 2005). The argument is that at the macro level both variables depend on the state of the economy (systematic risk). At the micro level the relationship between the two is derived from the demand and supply of defaulted securities. This argument suggests that, like default, recovery rate has a systematic risk component and should therefore attract a risk premium and sufficiently incorporated into credit risk assessments.

Second, the modelling of credit risk has been done mainly using large firms from advanced countries. The empirical literature on credit risk modelling among small scale industries and in the developing world on the subject is rather sparse. There are several reasons why credit risk may not be uniform across firms of different sizes operating in different environments. In the first place, unlike small scale enterprises, large corporations can issue bonds and obtain other sources of funding such as bank loans to manage their funding difficulties and credit risk. Also, small scale enterprises often either lack collateral or possess collateral of low quality which exacerbates their recovery difficulties. Thirdly, the differences in ownership characteristics between corporates and small scale enterprises yield different agency costs which could result in different credit risk outcomes. Lastly, the institutional, political and economic calculus differ between advanced and developing economies and are sufficient reasons why studies done in developed country contexts would be far removed from developing country realities.

Another problem with the extant literature on credit risk modelling of recovery rates is the dominance of practitioner literature. The scanty academic literature on recovery rates are either case studies or lump bonds and loans together in a unitary sample (Khieu, Mullineaux & Yi, 2012). Khieu et al. (2012) argue that given the differences between loans and bonds in terms of securitisation (bank loans are more secured than bonds), it should be expected that different factors determine recovery rates in the two market segments. In the scheme of capital structure, bank loans have a first claim to the liquidation value of the firm before bondholders. This implies that defaults on bank loans will be lower than on bonds. This study focuses on loans since small scale enterprises in developing countries mostly do not even have the capacity to issue bonds.

This study makes a two-fold contribution to the literature. The first contribution is methodological, applying a two-step approach to analyse probability of default and rate of default. So far, no study has carried out analysis using this two-step procedure where loan repayment default is modelled in the first stage to determine the factors influencing probability of default and modelled in the second stage to determine factors influencing the quantum (rate) of default. This novel approach helps overcome endogeneity and reverse causality (both common) in one stage models. The one stage approach of either analysing probability of default or determinants of rate of default imposes an assumption which would be explored in the literature review section. Second, this study joins the few empirical studies (e.g., Kosak & Poljsak, 2010; McCann & McIndoe-Calder, 2012; Menton & Sherman, 2014) that apply recovery rate modelling to SSEs. However, this study departs significantly from these studies in the methodology employed as afore-mentioned. Most studies (see Awunyo-Vitor, 2012; Wongnaa and Awunyo-Vitor, 2013; Anigbogu, Onuga, Onyeugbo & Okoli, 2014) also concentrate on individual or household characteristics other than enterprise characteristics. This study uses both enterprise level characteristics and household level characteristics.

The rest of the paper is structured as follows. The review of related literature immediately follows. After the literature review, the methodology is presented next. Presentation and discussion of results follows the methodology. Finally, conclusion and policy implications are presented.

Literature Review

There is considerable literature on credit risks of financial institutions. In general, a review and summary of the numerous credit risk studies show that studies related to recovery rate are evolving unlike those of probability of default where extensive work has been done (see Grunert & Weber, 2009; Khieu et al., 2012; Bastos, 2010; Bade, Rosch & Scheule, 2011). For example, Bastos (2010) notes that the recent attention on

recovery rates was triggered by the advent of the new Basel Capital Accord. Also, most of the studies in recovery rates are skewed towards bond recovery rates rather than loan recovery (Khieu et al., 2012; Bastos, 2010; Dermine & Neto de Carvalho, 2006).

On the theoretical front, there are two schools of thought on the relationship between probability of default and recovery rate. While Frye (2000), Altman et al. (2005), and Bruche and González-Aguado (2010) argue that probability of default and recovery rate are correlated, Grunert and Weber (2009), and Jarrow, Lando and Turnbull (1997) show that the two are uncorrelated. The theorists of correlated probability of default and recovery rates posit a negative correlation between the two. The implication of correlation is that, modelling probability of default and recovery rate as standalone models would yield biased estimates (Bade et al., 2011; Bruche & González-Aguado, 2010). Another notable issue in the literature is the rigorous application of traditional multivariate statistical and econometric models (Lin, 2009). The prominent ones include multivariate discriminant analysis, logit, probit, non-parametric models such as mathematical programming, classification trees and neural networks.

In terms of the empirics, Demirgüç-Kunt and Detragiache (1998) and Bonfim (2009) explore a link between banking crisis and macroeconomic developments of countries. They observe that banking crisis is more likely to occur in a weak macroeconomic environment. Authors such as Pederzoli and Torricelli (2005), and Jimenez and Saurina (2004), offer explanations supporting the findings of Demirgüç-Kunt and Detragiache (1998) and Bonfim (2009). They note that high default rates during recessions are consequences of unsustainable and risky transactions pursued during periods of expansions. In contrast to these macroeconomic-based findings, the results of Johnson, Boone, Breach and Friedman (2000) show the superiority of corporate governance over macroeconomic measures as a predictor of banking crisis. They observe that in countries with weak corporate governance, financial distress problems are more likely to emerge.

Methodologically, two options have been identified in the loan repayment default literature. The first option is where repayment default is modelled as a discrete outcome. That is, either a client defaults or not. In these studies, categorical (binary) choice models such as probit and logit are employed for analyses (see, for instance, Awunyo-Vitor, 2012; Mokhtar, Natea & Gan, 2012). The second option is where repayment or default is modelled as a continuous outcome. That is, the amount of repayment default (given as the difference between amount of loan, including interest, and the amount paid within the stipulated time). In these cases, multiple regression, test of means, chi-square, Tobit and discriminant analysis are used (see, for instance, Sharma & Zeller, 1997; Eze & Ibekwe, 2007; Oladeebo & Oladeebo, 2008; Munene & Guyo, 2013; Anigbogu et al., 2014).

In using these approaches, which can be regarded as one step approaches, one methodological issue arises. The assumption that the same set of factors determines probability of repayment or default on one hand and amount of repayment or default on the other is implied. The problem with this assumption is that, it is possible that a variable could have different effects on the probability and rate of default. The assumption is restrictive and hence, makes the one step approach restrictive. This study proposes the use of a two-stage approach as a means of overcoming the restrictiveness of the one step approach. The benefit of using this two-step procedure thus, is the ability to account for the factors influencing the probability of default and the amount of default separately but in a simultaneous manner.

Analytical Method

Empirical Model and Choice of Econometric Model

In modelling repayment default with a two-step approach, two dependent variables are identified. Whether a loan beneficiary was able to fully pay back the loan or not represents the dependent variable for the first stage model and the rate of default if the beneficiary was not able to fully pay back represents the dependent variable for the second stage model. Thus, the first step is the usual approach in the literature where this study estimates the probability of loan default using the discrete approach. The second step is estimating the factors that determine the level (amount) of loan default using the continuous approach.

Given a typical small scale entrepreneur who benefitted from microcredit, P_i , for a period of time, T_i , at an interest rate, R_i , the simple interest payable at the end of the period would be expressed as:

$$I_i = \frac{P_i \times T_i \times R_i}{100} \quad (1)$$

Total amount to be repaid, A_i is the sum of loan amount, P_i and interest, I_i . At the end of the loan period, T_i , with amount of repayment, A_i^* , the following mutually exclusive scenarios can occur:

$$A_i^* = A_i | T_i \quad (2a)$$

$$A_i^* < A_i | T_i < (T_i + t) \quad (2b)$$

where t is a period of at least 1 day after the last stated repayment date. This implies that the operational definition of default is a loan that was delinquent for at least 1 day after the stated duration. It should be noted that within the period of the loan, a client could default in a particular day, week, fortnight or month depending on the repayment

schedule. This default is not what is defined here. It is the default at the end of the entire loan duration which is defined.

The rate of repayment, RR_i from these scenarios would be defined by:

$$RR_i = \frac{A_i^*}{A_i} \times 100 \tag{3}$$

The default rate, DR_i from scenario (2a) would be defined by:

$$DR_i = 100 - RR_i = 0 \tag{4}$$

While the default rate in scenario (2b) will be defined by:

$$DR_i = 100 - RR_i = \omega \tag{5}$$

where ω is nonnegative and not zero ($\omega > 0$).

Equations (2a) and (2b) define the discrete dependent variable in the first stage and equation (5) defines the continuous dependent variable in the second stage. If a single variable for (2a) and (2b), ψ , is defined, then, the study seeks to model:

$$P(\psi_i = 1) = P(\omega_i > 0) = X_i\alpha + \varepsilon_i \tag{6}$$

$$\omega_i = Z_i\beta + \mu_i \tag{7}$$

where equations (6) and (7) are loan default and rate of default models respectively; ψ_i assumes the value of 1 if an individual defaults and 0 otherwise. ω_i is rate of default and X_i and Z_i are the vectors of explanatory variables hypothesised to influence probability of loan default and rate of default respectively. These variables are explained in the next section. α and β are the parameters to be estimated and $\varepsilon_i \sim N(0,1)$ and $\mu_i \sim N(0, \sigma_\mu^2)$.

Departing from these previous studies into the realm of two-step approach, the model to tackle the issue under consideration also changes. An important issue worthy of mention is the fact that rate of default is only observed for a subset of the sampled population (those with scenario of equation (2b)) because enterprises with scenario (2a) would have zero rates of default. This introduces selection bias as indicated by Heckman (1979). As a consequence, estimates of equations (6) and (7) based on least squares would be biased and inconsistent (Greene, 2003). There are a number of alternatives to address the selectivity bias and provide unbiased, consistent and efficient estimates. The first is the widely used Tobit model (proposed by Tobin, 1958) developed to alleviate the problems caused by least squares. Though the Tobit model is implementable in this paper, its imminent restriction (the assumption of variables which determine the probability of default also determining the rate of default) makes it unusable. In other

words, Tobit would only estimate equation (7) and assume that what determines (7) determines (6) with the same magnitude and direction.

The second alternative is to employ the Heckman sample selection model which estimates (6) with probit and (7) with least squares after including among the covariates the inverse mills ratio (IMR) derived from the probit to account for selection bias. Following from the observation of Ricker-Gilbert, Jayne and Chirwa (2011), the Heckman model is designed for incidental truncation, where the zeros are unobserved values. However, in this paper a zero value in the data would reflect an enterprise' optimal choice rather than a missing value. It would be erroneous to equate these missing observations to zero (Olwande & Mathenge, 2012). Therefore, the Heckman model is not also used.

The third alternative is the Cragg's double hurdle model (proposed by Cragg, 1971). It is a corner solution model just like the Tobit. The fact that the Tobit and the double hurdle models are corner solution models makes them appropriate to this study than the Heckman model. Considering the restrictiveness of the Tobit model, the double hurdle model is the feasible one adopted. Following Burke (2009), the double hurdle model can be specified theoretically as:

$$f(\psi, \omega | X_i, Z_i) = \{1 - \Phi(X_i \alpha)\}^{1(\psi=0)} \left[\Phi(X_i \alpha) (2\pi)^{-\frac{1}{2}} \sigma^{-1} \exp\left\{-\frac{(\omega - Z_i \beta)^2}{2\sigma^2}\right\} / \Phi(Z_i \beta / \sigma) \right]^{1(\psi=1)} \quad (8)$$

where Φ is the standard normal cumulative distribution function and exponential indicator functions are $1(\psi=0)$ and $1(\psi=1)$. The rest of the variables are as defined before. Equation (8) integrates equations (6) and (7) as one model and specifies a probit model to determine the probability of default and the truncated normal model for rate of default simultaneously using Burke's (2009) *craggit* command in Stata. However, the parameters of the first and second hurdles can also be estimated using probit and truncated normal regression respectively. The double hurdle model is a maximum likelihood estimator. The fundamental assumption underlying the model is that the errors of the two hurdles are independent, normally distributed and with zero covariance. One advantage of the double hurdle is its ability to nest the Tobit model thus allowing for the testing of the restrictions implied by the Tobit hypothesis (Burke, 2009; Olwande & Mathenge, 2012).

Testing and Controlling for Endogeneity

The choice of the double hurdle model does not in itself guarantee unbiased, consistent and efficient estimates if other issues are not taken care of. It is suspected that the amount of loan and previous profit is potentially endogenous to the probability and rate of default. In as much as several factors cause enterprises to default, the amount of

loan is expected to be the prime cause of default and the quantum of the default. High interest rates are usually seen as a constraint to loan repayment, but even such interest rates work through the amount of loan. Also, from the microfinance institution's (MFI's) standpoint, the amount of loan given out would depend on the ability to repay which is usually referred to as creditworthiness. The implication of this suspicion of amount of loan and profit being endogenous is that there may be correlation between the error terms in a reduced form equation of loan amount and profit, and in equations (6) and (7) respectively.

To resolve this, the control function (CF) approach is adopted. Generally, the CF involves estimating a reduced form function with the suspected endogenous variable(s) (loan amount and profit in this case) as the dependent variable(s) and estimating the residual(s) of this function for inclusion as covariate(s) in models (6) and (7). The significance of the coefficient(s) on the residual(s) both tests and controls for endogeneity. However, standard errors associated with the estimated residuals in the double hurdle model are no longer valid. We overcome this by bootstrapping standard errors for the residuals. The loan amount and profit reduced form models are estimated using ordinary least squares (OLS).

One requirement for the use of the CF approach is the introduction of an instrumental variable in the loan amount and profit reduced form models that is not in the probability and the rate of default models. It is expected that the instrumental variable should be correlated with the potentially endogenous variables (loan amount and profit) but not correlated with the probability and the rate of default variables. The instrumental variable considered is type of enterprise engaged in [ETYPE] which is believed to have an influence on loan amount and profit and at the same time has minimal chances of correlating with loan default and rate of default. The implication is that ETYPE is assumed to be exogenous in the loan default and rate of default models.

Study Area, Data and Variables

The Upper West Region has eleven administrative districts. The study was conducted in the Wa Municipality of the Upper West Region which is the only municipality of the eleven districts. The population of the municipality stands at 107,214 representing 15.3% of the regional total population. It is therefore the largest in terms of population. Out of the population figure, 71,051 and 36,163 are urban and rural dwellers respectively.

Primary data were collected from a survey of 200 small scale entrepreneurs using structured questionnaires. The survey was conducted between May and July, 2014. The first process of the data collection was the listing of beneficiaries of microcredit from microfinance institutions who operated SSEs in the municipality. The target

population thus comprised beneficiaries who had completed the entire period within which the loan was supposed to be paid back. At the time of data collection, some of the target population were on fresh loans. These beneficiaries were not part of the target population based on the current loans but on the previous loans that had their repayment period elapsing. In the process of listing, 497 SSEs were recorded. The second process was the application of simple random sampling to the list generated to draw respondents. Due to logistical constraints, 200 SSEs (representing about 40% of the 497 SSEs) were taken as the sample size.

Variables considered are categorised as borrower, enterprise and loan/lender characteristics. Borrower characteristics include gender of borrower, age of borrower, education of borrower and number of dependents. Gender [GEN] is measured as a dummy taking on a value of 1 if a client is male and 0 otherwise. This could have a positive or negative effect on probability and rate of default. Age [AGE] is a continuous variable measured by number of years of the borrower. This could have a positive or negative effect on probability and rate of default. Educational level [EDU] is a continuous variable measured by number of completed years of schooling and is expected to decrease probability and rate of default. Number of dependents [DEPN] is measured by the number of people under the care of the client. The expectation is not certain. Enterprise characteristics include age of enterprise, size of enterprise, level of previous profit before the loan and type of enterprise. Age of enterprise [EAGE] measures the number of years of existence of a client's enterprise up to the year of the loan. This is expected to decrease probability and rate of default. Size of enterprise [ESIZE] measures the number of people employed to work in the enterprise in the year of the loan and is expected to decrease probability and rate of default.

Profit [PRF] measures profit earned or loss incurred in the year preceding the year of loan and is expected to decrease probability and rate of default. Type of enterprise operated [ETYPE] is a dummy taking a value of 1 if enterprise is in trading of agricultural commodities and 0 otherwise. Given that agriculture is the mainstay of the region, the variable seeks to explore how trading in unprocessed agricultural commodities affect amount of loan and profit. It is expected to increase loan amount and profit. Loan characteristics include interest rate, loan period, repayment schedule and loan amount. Interest rate [INTR] measures the rate of interest per annum on loan. This is expected to increase probability and rate of default. Loan period [LNPRD] is the duration of loan measured in years and is expected to decrease probability and rate of default. Repayment schedule [RESCH] is a dummy variable taking the value of 1 if repayment is made fortnightly and 0 otherwise. Its expectation is not certain. Loan amount [LNAMT] is the total amount of money taken as loan and is expected to increase probability and rate of default.

Empirical Results and Discussions

Socio-economic Characteristics of Respondents

Socioeconomic characteristics and other descriptive statistics of respondents are presented in Table 1. The mean age of respondents is 38 years. About 54% is male while 65% is married. Educational status showed that on average respondents spent about 9 years in school. Average number of dependents of respondents is 4 people. Mean age of enterprise is about 9 years while the mean size of enterprise is 5 people. Enterprises made an average profit of GH¢108.26 in the preceding year to the loan. Majority (51%) of respondents paid back their loans fortnightly. This was followed by those who paid on monthly basis (32.5%). Only 1% had a schedule of daily payments while 15.5% had a schedule of weekly repayments.

Interest rate averaged 23.3% per annum. This ranged from 11% to 38% per annum. Duration of loan repayment ranged from 1 year to 4 years with a mean of about 2 years. Agro-processing represents the dominant enterprise run by respondents as 53% are engaged in it. This is followed by enterprises in the trading of unprocessed agricultural commodities (41.5%). Store operating enterprises (sales of provisions) represent the least enterprise (5.5%) engaged in by respondents. The average amount of loan accessed was GH¢410.13. Loan amount due ranged from GH¢90 to GH¢1750 with an average of GH¢608.61. Amount of loan paid back averaged GH¢570.11 with a minimum of GH¢90 and a maximum of GH¢1750.

At the end of the repayment period, 34.5% of respondents defaulted in paying back their loans. The rate of repayment ranged from 54.1% to 100% with a mean repayment rate of about 94%. On the other hand, rate of default had a minimum of 3.4% and a maximum of 45.9% with a mean of 18.3%.

Table 1: Socio-economic characteristics and descriptive statistics of variables

Discrete variables		Continuous variables				
Variable	Freq. (percent)	Variable	Mean	Std Dev.	Min	Max
Loan status		Loan amount	410.13	240.42	75	1000
Repaid fully	131(65.5)					
Defaulted	69(34.5)					
Gender:		Loan amount due	608.61	376.29	90	1750
Male	107(53.5)					
Female	93(46.5)					
Marital status:		Loan amount paid	570.11	363.87	90	1750
Married	130(65.0)					
Never married	43(21.5)					
Widow(er)	19(9.5)					
Divorced	8(4.0)					
Education:		Rate of repayment	93.69	10.94	54.1	100
None	39(19.5)					
Primary	26(13.0)	Rate of default	18.3	11.31	3.38	45.9
Middle/JHS	62(31.0)					
SHS	28(14.0)	Age	37.74	9.06	21	60
Diploma	24(12.0)					
Undergraduate	19(9.5)					
Postgraduate	2(1.0)					
Enterprise type:		Education	8.58	5.40	0	20
Agro-processing	106(53.0)					
Agric. trading	83(41.5)	Number of dependants	4.04	2.75	0	15
Provision store	11(5.5)					
Repayment schedule:		Age of enterprise	8.51	5.11	1	25
Daily	2(1.0)					
Weekly	31(15.5)					
Fortnightly	102(51.0)	Size of enterprise	4.5	2.52	1	16
Monthly	65(32.5)					
Reason for default:		Profit	108.26	306.89	-750	850
Wrong investing	16(23.2)	Interest rate	23.31	5.14	11	38
Diversion	13(18.8)	Loan period	1.94	0.94	1	4
Unfavourable repay. schedule	22(31.9)					
Insufficient loan	9(13.0)					
Natural calamity	9(13.0)					

From the 69 defaulters, 31.9% attributed their default to difficult or unfavourable repayment schedule, 23.2% to wrong investment decisions, 18.8% to diversion of loan and 13% to insufficient loan and natural calamities apiece.

Determinants of Loan Default and Rate of Default

Table A1 (see appendix) presents the results of the reduced form models of loan amount and profit. First of all, the instrument (ETYPE) is highly relevant in both reduced form models. As a rule of thumb in the case of a single instrument, the t-value of the instrument should be greater than 3.2 or the corresponding p-value below 0.0016. The t-values for the instrument in the loan amount and the profit models are respectively 5.31 and 3.59 with p-value of 0.000 each. For test of exogeneity (see Table A2 in appendix), we regressed the instrument on the residuals of loan amount and profit separately. The results show that the instrument bears no statistically significant relationship with the residuals. This implies that type of enterprise is truly exogenous in relation to the probability of default.

The estimates of the probability of loan repayment default and rate of default are presented in Table 2. Table 2 has the probability and rate of default model estimates through the step-by-step approach and the double hurdle approach. The results show that, the standalone probit results and the double hurdle probit results are the same. However, the standalone truncated regression and that of the double hurdle are different in the parameters of all the variables except for EDU which, in addition to the parameter, the signs also differ. Also, while AGE and DEPN are significant determinants in the standalone truncated regression, they are not significant in the double hurdle estimates. These strengthen the argument of bias of the one-step approach.

Table 2: Estimates of determinants of loan repayment default and rate of default

Variable	Step-by-step estimates				Double hurdle estimates			
	Probit Regression		Truncated Normal Regression		Hurdle 1 Probability of Loan Default (Probit Regression)		Hurdle 2 Rate of Default (Truncated Normal Regression)	
	Coefficient	Robust std. error	Coefficient	Robust std. error	Coefficient	Robust std. error	Coefficient	Robust std. error
GEN	1.2001 [*]	0.6699	-0.3590	1.0037	1.2001 [*]	0.6699	-2.3871	2.3806
AGE	0.0214	0.0561	0.1834 ^{***}	0.0570	0.0214	0.0561	0.1041	0.1224
EDU	-0.2663 ^{**}	0.1058	-0.0392	0.0900	-0.2663 ^{**}	0.1058	0.0980	0.1829
DEPN	0.5959 ^{***}	0.1821	-0.3585 ^{**}	0.1509	0.5959 ^{***}	0.1821	-0.5161	0.4559
EAGE	-0.0192	0.1575	-0.5234 ^{***}	0.0928	-0.0192	0.1575	-0.7564 ^{**}	0.3585
ESIZE	-1.2957 ^{***}	0.2482	-0.8218 ^{***}	0.1914	-1.2957 ^{***}	0.2482	-2.1830 ^{***}	0.7302
INTR	0.1852 ^{***}	0.0712	0.9732 ^{***}	0.1494	0.1852 ^{***}	0.0712	2.1069 ^{***}	0.2229
LNPRD	-3.1848 ^{***}	1.1678	-1.2572 ^{**}	0.4938	-3.1848 ^{***}	1.1678	-3.3501 ^{***}	1.2556
RESCH	-2.5210 ^{***}	0.8859	-0.5551	0.9107	-2.5210 ^{***}	0.8859	-1.3912	1.8483
PRF	-0.0705 ^{***}	0.0214	-0.0117 ^{***}	0.0022	-0.0705 ^{***}	0.0214	-0.0095 ^{***}	0.0037
LNAMT	0.0609 ^{***}	0.0199	0.0064 ^{***}	0.0024	0.0609 ^{***}	0.0199	0.0184 ^{***}	0.0050
LNAMT_res	-0.0544 ^{***}	0.0187			-0.0544 ^{***}	0.0062		
PRF_res	0.0579 ^{***}	0.0179			0.0579 ^{***}	0.0050		
CONSTANT	-13.2149 ^{***}	4.4221	-11.7801 ^{***}	3.7451	-13.2149 ^{***}	4.4221	-31.6884 ^{***}	6.8065
Sigma			6.2950 ^{***} 0.4342		6.8766 ^{***} 0.7714			
No. of observations	200		69		200 69			
Wald chiz	84.81 ^{***}		242.42 ^{***}		84.81 ^{***}			
Log pseudo likelihood	-10.5469		-651.7376		-230.5690			

^{***} $p < 0.01$, ^{**} $p < 0.05$ and ^{*} $p < 0.10$. The standard errors of the residuals (LNAMT_res and PRF_res) in the double hurdle results are bootstrapped at 150 repetitions since the normal standard errors from the estimates are not valid.

Source: Regression estimates from survey data, 2014

Generally, the results of the double hurdle model are robust compared with those from the step-by-step approach. The discussions of the results are based on the double hurdle results since the weakness of the step-by-step approach has been underscored.

The loan amount and profit residuals (LNAMT_res and PRF_res) from the reduced form models are statistically significant at 1% implying that loan amount and profit are

endogenous in the loan default model. They are not significant in hurdle 2 hence they are omitted in the final estimation. The estimates in hurdle 1 show that probability of loan repayment default is influenced by gender of client, education level, number of dependants, size of enterprise, rate of interest, duration of loan, repayment schedule, level of profit and loan amount. Rate of default on the other hand is influenced by age of enterprise, size of enterprise, rate of interest, duration of loan, level of profit and loan amount.

It is found that males are 120% more likely to default. Men are generally more risk loving than women and for that matter can apply loans in very risky ventures which may end up failing. Another explanation for this result is that, women are generally more skilful at agro-processing and trading and for that matter are less likely to default when borrowed funds are applied to these enterprises. This finding also, confirms the observations by Abu et al. (2016) that women are relatively more industrious than men in the operation of SMEs. Gender does not however determine rate of default.

This confirms the observation of Abu et al. (2016) that women are seen as relatively more industrious than men in the operation of SMEs. Trading and agro-processing are activities that women are comparatively skilful at and thus may be the reason for this finding. Gender does not however determine rate of default.

More educated clients are 26.6% less likely to default. This is consistent with the work of Bridges and Disney (2004), and Eze and Ibekwe (2007) who observed that borrowers with low level of education have higher risks of debt. The likelihood of managerial differences between more educated clients and less educated ones might be responsible for this observation. More educated borrowers might have a better economic standing and able to make more advantageous financial decisions (Schicks, 2014). Level of education however, does not influence rate of default.

Entrepreneurs with more dependents are 59.6% more likely to default. This is consistent with a priori expectation. Large dependents may serve as a strain on business expansion and growth due to huge consumption expenditures. Thus, huge expenditures stifle prompt repayment. Number of dependents does not determine rate of default.

While age of enterprise does not determine probability of default, it is a negative determinant of rate of default. Enterprises that have operated for relatively longer years record smaller default rates as compared to young enterprises. Specifically, they are 0.76% less likely to record high default rates. This confirms the finding of Bastos (2010) that older firms exhibit better recoveries. This finding endorses the role of experience in the prudent management of loans. The findings suggest that even if experienced enterprises default, they leverage their vast experience to reduce the rate of default.

Enterprises with larger number of workers are 129.6% less probable to default in loan repayments. This meets a priori expectation. Size is an indicator of the economic muscle of an enterprise. Kiboki et al. (2014) view size of enterprise as a performance indicator. The argument therefore is that large enterprises have the capacity to efficiently utilise loans to derive significant benefits that make them more likely to fully pay back their loans than small enterprises. Size of enterprise also negatively influences rate of loan default. This implies that enterprises with larger number of workers are 2.2% less likely to default with high rates.

The positive estimate of interest rate shows that enterprises that secured loans with higher interest rates are more probable of defaulting. Such enterprises are 18.5% more likely to default. This observation stresses the role of high interest rate in loan defaults and advances the arguments of the effect of high interest rate in stifling enterprise development as well as the sustainability of credit institutions. With respect to rate of default, interest rate correlates positively with rate of loan default. Clients who had higher interest rates are 2.1% more likely to record higher rates. This further indicates the effect of interest rate on the quantum of default.

It is observed that enterprises that paid back their loans over longer period of time are 318.5% less likely to default. This is highly anticipated since spreading repayment over long time periods offers enterprises the opportunity of spreading loan burden in a manner that keeps pace with the cash flow of the enterprise. This observation agrees with Jimenez and Saurina (2004) who argue that loans with longer durations have lower probability of defaults due to the fact that credit institutions do proper and thorough examination of the application since the borrower's financial health could change significantly over such a long period. However, Eze and Ibekwe (2007) find that borrowers with long durations record more outstanding loan amount. On rate of default, duration of loan is estimated as a negative determinant. This means that enterprises that paid back their loans over longer period of time record small rates of loan default. Such enterprises are 3.4% less likely to record high rates.

The negative estimate of repayment schedule shows that enterprises that make repayment fortnightly are 252.1% less likely to default as compared to those who repay daily, weekly and monthly. A plausible reason could be that repayment in fortnight is an optimal time that favours efficient loan repayment. Probably, daily and weekly repayments are too short a time while monthly repayments are too long a time to ensure prompt repayments. Daily and weekly schedules are probably strenuous due to inadequate cash flow while monthly schedules are problematic due to household expenditures. The observation implies that cash flow of enterprises correlate more with fortnight repayments than the other repayment periods. Repayment schedule does not however influence rate of default.

It is found that enterprises that make large profits are 7.1% less likely to default in loan repayment. The result is consistent with those of Bonfim (2009), Wongnaa and Awunyo-Vitor (2013), and Oladeebo and Oladeebo (2008). This is highly anticipated since loan repayments principally are made from enterprise profits. Therefore, the larger the profit the more probable an enterprise is to fully pay loans without defaulting. Also, profit is another indicator of an enterprise's performance and an indicator of repayment ability. Hence, more profitable firms should have a more solid financial situation (Bonfim, 2009) to enable repayments. In terms of default rate, the level of profit negatively determines the rate of default which implies that higher profits induce lower default rates. The larger the profit, rate of default is 0.01% smaller.

Amount of loan is a positive determinant of probability of loan repayment default. This means that enterprises that secured larger loan amounts are 6.1% more likely to default. This is expected since a large amount of loan represents a large cost and responsibility for an enterprise to meet. This contradicts the finding of Awunyo-Vitor (2012) and Jimenez and Saurina (2004) but confirms the finding of Sharma and Zeller (1997), and Eze and Ibekwe (2007). With respect to rate of default, amount of loan is a positive determinant. This means that enterprises that secured large amounts of loans record large rates of default. Such enterprises record 0.02% more rate of default. Just as such enterprises are more likely to default, they default with large rates. This corroborates the finding of Dermine and Neto de Carvalho (2006) and Bastos (2010) who observed that size of loan negatively influences recovery rate.

Conclusion and Policy Implications

The success stories of MFIs in providing credit to SSEs depend on repayment performance. Repayment default is a fundamental bane, crippling and stifling the efforts of these microcredit granting institutions. This paper sought to apply a two stage approach to analysing probability and magnitude of loan default as a way of avoiding the assumption imposed by various studies that, the same factors determine the probability and magnitude of default. The control function approach was used to test and control for potential endogeneity of loan amount and profit in relation to probability of default. In summary, the results showed that enterprise size, interest rate, loan duration, level of profit and loan amount are the simultaneous determinants of probability and rate of default. However, the age and educational attainment of the enterprise owner, number of dependents and loan repayment schedule influence the probability of default but not the rate of default.

These findings imply that microfinance institutions must be conversant with the factors that influence the probability of defaults separate from those that influence the rate

of default. Thus, they would require separate measures to reduce the probability of default and the rate of default. From policy perspective, the National Board for Small Scale Industries (NBSSI) should intensify training programmes aimed at enhancing the managerial and technical capabilities of entrepreneurs to ensure healthy and productive management as this would promote enterprise growth and increase profit margins and hence reduce defaults. One way to make training programmes proactive is to design industry specific programme modules. Also, microfinance institutions should consider reviewing repayment periods to suit the peculiarities of each enterprise in order to diffuse the burden of repayment. While large loan amounts are problematic, appropriate targeting of large enterprises for large amounts could reduce defaults. Furthermore, government's role in ensuring a stable macroeconomic environment that favours moderate interest rates cannot be overemphasised.

References

- Abu, B.M., Domanban, P.B. and Sekyi, S.** (2016). Credit market participation by women-owned small scale enterprises in Wa and Jirapa districts of the Upper West region of Ghana. *Ghanaian Journal of Economics*, 4(2016), pp. 71-97.
- Altman, E.I., Brady, B., Resti, A. and Sironi, A.** (2005). The link between default and recovery rates: theory, empirical evidence and implications. *Journal of Business*, 78, pp. 2203-27.
- Altman, E.I., Resti, A. and Sironi, A.** (2004). Default recovery rates in credit risk modelling: a review of the literature and empirical evidence. *Economic Notes*, 33(2), pp. 183-208.
- Anigbogu, T.U., Onugu, C.U., Onyeugbo, B.N. and Okoli, M.I.** (2014). Determinants of loan repayment among cooperative farmers in Awka North L.G.A of Anambra State, Nigeria. *European Scientific Journal*, 10(22), pp. 168-190.
- Awunyo-Vitor, D.** (2012). Determinants of loan repayment default among farmers in Brong Ahafo Region of Ghana. *Journal of Development and Agricultural Economics*, 4(13), pp. 339-445.
- Bade, B., Rosch, D. and Scheule, H.** (2011). Default and recovery risk dependencies in a simple credit risk model. *European Financial Management*, 17(1), pp. 120-144.
- Bastos, J.A.** (2010). Forecasting bank loans loss-given-default. *Journal of Banking & Finance*, 34, pp. 2510-2517.
- Bonfim, D.** (2009). Credit risk drivers: evaluating the contribution of firm level information and of macroeconomic dynamics. *Journal of Banking & Finance*, 33, pp. 281-299.

- Bridges, S. and Disney, R.** (2004). Use of credit and arrears on debt among low-income families in the United Kingdom. *Fiscal Studies*, 25(1), pp. 1–25.
- Bruche, M. and González-Aguado, C.** (2010). Recovery rates, default probabilities, and the credit cycle. *Journal of Banking & Finance*, 34, pp. 754–764.
- Burke, W.J.** (2009). Fitting and interpreting Cragg's tobit alternative using Stata. *Stata Journal*, 9(4), pp. 584–592.
- Cragg, J.G.** (1971). Some statistical models for limited dependent variables with application to the demand for durable goods. *Econometrica*, 39(5), pp. 829–844.
- Demirgüç-Kunt, A. and Detragiache, E.** (1998). The determinants of banking crises: evidence from industrial and developing countries. Working Paper 1828, March, IMF.
- Dermine, J. and Neto de Carvalho, C.** (2006). Bank loan losses-given-default: a case study. *Journal of Banking & Finance*, 30, pp. 1219–1243.
- Eze, C.C. and Ibekwe, U.C.** (2007). Determinants of loan repayment under the indigenous financial system in Southeast, Nigeria. *The Social Sciences*, 2(2), pp. 116–120.
- Frye, J.** (2000). Collateral damage risk. April, pp. 91–94.
- Greene, W.H.** (2003). *Econometric Analysis*. 5th edition. New Jersey: Pearson Education International.
- Grunert, J. and Weber, M.** (2009). Recovery rates of commercial lending: empirical evidence for German companies. *Journal of Banking & Finance*, 33, pp. 505–513.
- Heckman, J.** (1979). Sample selection bias as a specification error. *Econometrica*, 47(1), pp. 153–161.
- Jarrow, R., Lando, D. and Turnbull, S.** (1997). A markov model for the term structure of credit risk spreads. *Review of Financial Studies*, 10, pp. 481–523.
- Jarrow, R.A.** (2001). Default parameter estimation using market prices. *Financial Analysts Journal*, 57(5), pp. 75–92.
- Jimenez, G. and Saurina, J.** (2004). Collateral, type of lender and relationship banking as determinants of credit risk, Working Paper, Banco de Espana.
- Johnson, S., Bone, P., Breach, A. and Friedman, E.** (2000). Corporate governance in the Asian financial crisis, 1997–98. *Journal of Financial Economics*, 58, pp. 141–186.
- Khieu, H.D., Mullineaux, D.J. and Yi, H.-C.** (2012). The determinants of bank loan recovery rates. *Journal of Banking & Finance*, 36(4), pp. 923–933.

- Kiboki, A.K., Sakwa, M. and Kiriago, A.N.** (2014). The Relationship between small scale enterprises performance and access to credit from microfinance institutions in Mount Elgon Constituency, Kenya. *International Journal of Academic Research in Business and Social Sciences*, 4(9), pp. 507-518.
- Kosak, M. and Poljsak, J.** (2010). Loss given default determinants in a commercial bank lending: an emerging market case study. *Zb. rad. Ekon. fak. Rij*, 28(1), pp. 61-88.
- Lin, S.L.** (2009). A new two-stage hybrid approach of credit risk in banking industry. *Expert Systems with Applications*, 36, pp. 8333-8341.
- McCann, F. and McIndoe-Calder, T.** (2012). Determinants of SME loan default: the importance of borrower-level heterogeneity. Technical Research Paper, 06/RT/12, Central Bank of Ireland.
- McQuown, J.A.** (1997). Market versus accounting-based measures of default Risk. In *Option Embedded Bonds*, (Ed), Nelken, I. Chicago: Irwin Professional Publishing
- Menton, A. and Sherman, M.** (2014). Analysis of the Irish SME market using micro-data. Bulletin 39, IFC.
- Mokhtar, S.H., Nartea, G. and Gan, C.** (2012). Determinants of microcredit loans repayment problem among microfinance borrowers in Malaysia. *International Journal of Business and Social Research*, 2(7), pp. 33-45.
- Munene, H.N. and Guyo, S.H.** (2013). Factors influencing loan repayment default in microfinance institutions: the experience of Imenti North District, Kenya. *International Journal of Applied Science and Technology*, 3(3), pp. 80-84
- Oladeebo, J.O. and Oladeebo, O.E.** (2008). Determinants of loan repayment among smallholder farmers in Ogbomoso Agricultural Zone of Oyo State, Nigeria. *Journal of Social Sciences*, 17(1), pp. 59-62.
- Olwande, J. and Mathenge, M.** (2012). Market participation among poor rural households in Kenya. In Paper Presented at the International Association of Agricultural Economists Triennial Conference, Brazil. (18-24 August).
- Pederzoli, C. and Torricelli, C.** (2005). Capital requirements and business cycle regimes: forward-looking modelling of default probabilities. *Journal of Banking & Finance*, 29, pp. 3121-3140.
- Ricker-Gilbert, J., Jayne, T.S. and Chirwa, E.** (2011). Subsidies and crowding out: a double-hurdle model of fertilizer demand in Malawi. *American Journal of Agricultural Economics*, 93(1), pp. 26-42.
- Schicks, J.** (2014). Over-indebtedness in microfinance – an empirical analysis of related factors on the borrower level. *World Development*, 54, pp. 301-324.

- Sharma, M. and Zeller, M.** (1997). Repayment performance in group based credit programmes in Bangladesh. *World Development*, 25(10), pp. 1731-1742.
- Tobin, J.** (1958). Estimation of relationships for limited dependent variables. *Econometrica*, 26(1), pp. 24-36.
- Wilson, T.C.** (1998). Portfolio credit risk. *Economic Policy Review*, Federal Reserve Board of New York.
- Wongnaa, C.A. and Awunyo-Vitor, D.** (2013). Factors affecting loan repayment performance among yam farmers in the Sene District, Ghana. *Agris on-line Papers in Economics and Informatics*, V (2), pp. 111-122.