Pacific Rim Property Research Journal, 2015
Vol. 21, No. 3, 259–274, http://dx.doi.org/10.1080/14445921.2016.1140713

Impact of borrower's attributes on mortgage default: evidence from Nigerian lending market

T.B. Oyedokun*, A.O. Adewusi and M.O. Bello

Department of Estate Management, Federal University of Technology, Akure, Nigeria

(*Received 24 September 2015; accepted 8 January 2016*)

The need for proper identification of mortgage default factors has become a major focus of mortgage researches given the debilitating effect of default on mortgage market and real estate finance in particular. This paper therefore analyses the socio-economic attributes of borrowers as default triggers in residential mortgages of Primary Mortgage Institutions (PMIs) in Nigeria. Relevant data were collected on profiles of 305 borrowers randomly drawn from the credit databank of 36 PMIs. Using logistic regression (LR), payment-to-income ratio, type and sex of borrowers are found as significant mortgage default factors. With 68.2% overall prediction accuracy, LR is found appropriate for mortgage default prediction. However, the findings of this study also signal the complexity that is inherent employing socio-economic factors for default probability prediction.

Keywords: borrower; default; mortgage; payment-to-income; PMIs

Introduction

Housing plays critical roles in a nation's social welfare and national economic growth through employment provision and capital asset formation while also serving as useful index for measuring the living standard of people (Saka & Akin, 2007; Sanusi, 2003). Housing is generally accepted as the man's second most important need after food (Onyike, 2007). Its contribution to national gross domestic product is also significant being one of the largest employment generating sectors. The role of finance in housing development therefore cannot be over-emphasized. However, due to the ever increasing demand on public fund, housing finance has remained a burden too heavy for most governments to bear. Consequently, public housing provision has become almost an impossible endeavour especially in the developing countries such as Nigeria. Therefore, individuals in their quest to meet their demand for housing had resulted to the use of mortgage finance mostly secured from financial institutions such as the Primary Mortgage Institutions (PMIs) which were specially established for the purpose of providing mortgages for residential development (Nubi, 2000).

Ojo (2009) found in a survey that about two-third of the residential loan seekers prefer mortgages secured from the PMIs to loans from any other sources. This is a clear indication of the potential of mortgage finance in meeting the housing needs of the people. The mortgage lending as a process involves conveyance of interest in landed property by the owner (mortgagor) to the lender (mortgagee) for a certain

*Corresponding author. Email: tunbosunoyedokun@gmail.com

amount of money (loan) with the promise of repayment according to a specified amortization after which the property reverts to the owner without further encumbrance (Aluko, 2007; Ojo, 2009). According to Clayton (2007), the classical form of real estate debt finance is the mortgage, a loan secured by real property. The collateralized property is subjected to foreclosure clause if the borrower fails to meet up with the conditions of repayment as spelt out in the mortgage deed (Ojo, 2009). While the capability of PMIs to meet the required national housing finance need is subject to questions (Kama et al., 2013), default remains a major challenge to the lending process (Lin, Lee, & Chen, 2011; Okpara, 2009).

Generally, default occurs when a borrower breaches the mortgage conditions resulting into additional cost to the lenders. In relation to mortgage, default occurs when a borrower delays payment for between 30 and 90 days (Clauretie & Sirmans, 2003). Given the negative impact of default, researches have been preoccupied with developing frameworks for identifying mortgage default factors. While series of attempts have been made to evaluate the impact of borrower's characteristics on probability of default using regression models, such studies are still rare in the developing countries (Ajayi, 1992; Akinlo & Emmanuel, 2014; Hakim & Haddad, 1999; Lin et al., 2011; Somoye, 2010). This study therefore advances global knowledge on mortgage default drawing on evidence from a developing country. Delgadillo and Gallagher (2003) note that having a more comprehensive understanding of the role of borrower-related factors in default is beneficial for both practical and theoretical reasons. In the context of mortgage default literature in Nigeria, the current study is unique both in its methodology and the nature of sampled borrowers. The remaining part of this study is structured as follows. The next section reviews relevant empirical studies followed by the research methodology in the third section. Section four presents the data analysis and discussion of findings before the conclusion in the last section.

Literature review

The occurrence of mortgage default has been traced to a lot of factors including socio-economic characteristics of borrowers. Canner, Gabriel, and Woolley (1991) examined the effects of default risk and race in the analysis of mortgage lending. The study focused on the Federal Housing Authority and conventional loan markets of the United States (US). Adopting factor analysis, discriminant analysis, probit model and logit model, the study found that the probability of loan delinquency is positively correlated with household unemployment, receipt of government assistance, family size and marital status (divorced). Conversely, the study found that loan delinquency is inversely related to the age and household liquid asset holdings. The study pointed out on a more somber note that delinquency prediction consists of a large unexplained random component, as default can arise from events that are difficult to foresee. It implies that the use of ex-ante data has the ability to capture components that systematically affect delinquency and are observable to the lender at loan origination but ignores the more unpredictable ex-post components.

Munnell, Geoffrey, Lynn and McEneaney (1996) examined differential treatment affecting minorities' access to credit and opportunities for home-ownership in Boston's lending market using data from the Home Mortgage Disclosure Act and Panel Study of Income Dynamics. It was found that Black and Hispanic mortgage applicants were more likely to be rejected than their White counterparts indicating that the race of a borrower might be a significant determinant of default risk. Hakim and Haddad (1999) used a sample of 9000 conventional mortgages randomly selected from Freddie Mac portfolio to assess impact of borrower's attributes on default. The study found that default rate is positively correlated with a high loan-to-value (LTV) ratio and low-income borrowers with a large number of dependants. Conversely, default was found to be negatively correlated with size of loan, age of collateralized property and borrower's disposable income less of monthly financial obligations. Anderson and VanderHoff (1999) provide further evidence on the relationship between race and default in another US-based study. Their study employed probit model and found that Black households have higher marginal default rates than their White counterparts.

Lee and Liu (2001) note that borrowers' characteristics and family status are significant factors of mortgage default. Using data from the Directorate General of Budget, Accounting and Statistics, in the Executive Yuan in Taiwan, the study further found that borrower's payment-to-income (PTI) ratio and overall wealth have significant influence on default probability. The empirical results of Jacobson and Roszbach (2003) indicate that home ownership, annual income from wages, debt coverage ratio, age, sex, engagement in other loans are all significant factors of loan default. The study by Cairney and Boyle (2004) found that borrower's marital status and credit risk are significantly positively correlated while age, education and income exhibit significant inverse correlation with probability of default. Diaz-Serrano (2005) investigated the socio-economic determinants of mortgage delinquency across 12 European Union countries with special focus on income volatility. Using data retrieved from the European Community Household Panel over the period 1994–2001 and employing bivariate probit model, it was found that income volatility significantly increases mortgage delinquency even among borrowers with higher income profiles.

Furthermore, using data drawn from an inventory of active and foreclosed Federal Housing Administration homes in the state of Utah, US, Delgadillo and Gallagher (2003) found that race as a demographic attribute is a significant factor of home foreclosure. Given the fact that default is a necessary precursor to foreclosure, it implies that race is a potential determinant of mortgage default. Peter and Peter (2006) utilized data retrieved from the Australian Bureau of Statistics for an empirical assessment of default risk. Using logistic regression (LR) to model the effect of homeowners' socio-economic attributes and housing characteristics on probability of default, it was discovered that income, age of borrower, LTV ratio, marital status and educational status are all significant determinants of mortgage default. Conversely, employment status, number of dependants, family type, occupation, affordability indicator and location are found as insignificant factors of default.

Bandyopadhyay and Saha (2009) examined the factors driving demand and default risk in residential housing loans of Indian Housing Finance Institutions using data on 13,487 housing loan accounts approved in the period 1993–2007. Adopting LR, it was discovered that change in market value of mortgaged property vis-à-vis the loan amount and equated monthly instalment-to-income ratio are major factors of default. Age profile of borrowers, marital status, employment situation and house preference are also found to have significant effect on probability of default at different statistical significance levels. Lin et al. (2011) found that borrower's gender, job position, registered permanent residence, degree of relationship with guarantor, LTV ratio, status and regional location of collateralized property are directly correlated with default probability. Conversely, the study found that borrowers' education and loan size are inversely related with default probability. As can be deduced from the studies reviewed, characteristics of borrowers have significant impact on mortgage default occurrence.

In terms of application, regression model has been adopted in a number of studies and found to yield good result in predicting mortgage default (See Abdou & Pointon, 2011). Kung, Wu, Hsu, Lee, and Yang (2010) analyse the mortgage loans of five Taiwanese commerce banks to identify the key factors that influence prepayments and defaults. The study assesses the predictive power of LR using data on a total of 16,213 loans in five Taiwanese commerce banks over the period 2002–2007. The study reveals that as far as the overall predictability is concerned, LR is able to provide simplified results in the measurement of model variables concerning defaults and prepayments. LR is a statistical tool for making predictions when the dependent variable is dichotomous and the independent variables are continuous or categorical. Thus, LR extends the techniques of multiple regression analysis to situations in which the outcome variable is categorical (Kung et al., 2010). Furthermore, LR has been identified as a simple and appropriate technique for estimating the *log* of odds of default as a linear function of loan applicant's attributes (Bandyopadhyay & Saha, 2009). The odds ratio measures the probability that default will occur in the repayment of the mortgage. Another plus for LR is its flexibility in incorporating both the qualitative and quantitative factors in the estimation of mortgage default probability.

Data source and data collection

Data used for this study were obtained from the databanks of PMIs operating in Lagos metropolis. These are the financial institutions established through the Mortgage Institutions Decree No. 53 of 1989 to grant loans to individuals for the purchase, building, improvement or extension of a dwelling or commercial property (Ojo, 2009). Due to the need to preserve confidentiality, authors were not allowed direct access to the databases. Instead, the credit officers of the PMIs were provided with a specially designed bio-data spreadsheet to obtain the required information. The form provided spaces for entering the profiles of the borrowers except their names and other forms of identity. For the purpose of model development, the credit officers were informed to supply information on both defaulting and non-defaulting borrowers at reasonably equal percentages. It should be stated that securing data of this nature was utterly challenging due to existing confidential policies. This accounted for relatively low number of institutions that obliged to provide the required data. Eventually, the bio-data spreadsheets that were returned contained data on profiles of 305 borrowers all of which were found valid for the analysis.

While the authors acknowledge that the sample is small for the development of a high-precision logistic model, it is interesting to note that the findings of this study are not estranged to those of other studies (e.g. Hakim & Haddad, 1999; Lin et al., 2011) with seemingly larger samples. Besides, similar studies have safely used smaller samples. For instance, Leishman and Watkins (2004) assessed the decision-making behaviour of office occupiers based a survey of 119 occupiers using LR technique. Also, Afolabi (2010) analysed loan repayment among small scale farmers using regression model developed based on 286 loans. Using LR, Delgadillo and Gallagher (2003) assessed the impact of borrower-related factors on home foreclosure based on evidence from 179 cases. Although areas of application differ, the methods of analysis are the same. The authors are therefore confident that the current study is not odd in terms of its sample size.

Variables and model specification

Table 1 presents the definitions and measurements of the 14 variables used in the analysis including the dependent variable which is loan default status (LNDFT) measured as a dummy variable. The variable defines the category a borrower belongs with "1" representing those who have defaulted and "0" representing those who have not defaulted. The probability of default is then estimated using LR developed based on

Table 1. Operationalization of variables.

Variable code	Variable definition	Unit of measurement
Dependent variable		
LNDFT	Loan default status	1 if borrower defaults, 0 if otherwise
Independent variables		
LTV	Loan-to-value ratio	Actual value
DURA	Loan duration	Actual value in years
PTI	Payment-to-income ratio	Actual value
BTYPE	Type of borrower	1 if borrower is a first-time borrower, 0 if otherwise
HIST	Availability of borrower's credit history	1 if borrower's credit history is available, 0 if otherwise
EMP	Borrower's employment status divided into three categories: EMP1 to EMP3	
EMP1	Salary earner	1 if borrower is a salary earner, 0 if otherwise
EMP2	Self-employed	1 if borrower is self-employed, 0 if otherwise
EMP3	Retired/pensioner	1 if borrower is retired/pensioner, 0 if otherwise
BSEX	Borrower's sex	1 if male, 0 if otherwise
MSTAT	Borrower's marital status divided into three categories: MSTAT1 to MSTAT3	
MSTAT1	Married	1 if borrower is married, 0 if otherwise
MSTAT2	Single	1 if borrower is single, 0 if otherwise
MSTAT3	Widowed/divorced	1 if widowed/divorced, 0 if otherwise
AGE	Borrower's age at the date of loan origination	Actual value in years
EDU	Borrower's highest educational status	1 if borrower is a graduate, 0 if otherwise
DEPEND	Number of borrower's dependants as at date of loan origination	Actual value by headcount
GUARAN	Availability of guarantor	1 if guarantor is available, 0 if otherwise
TRIBE	Borrower's tribe divided into three categories: TRIBE1 to TRIBE3	
TRIBE1	Yoruba	1 if borrower is Yoruba, 0 if otherwise
TRIBE2	Igbo	1 if borrower is Igbo, 0 if otherwise
TRIBE3	Hausa	1 if borrower is Hausa, 0 if otherwise

explanatory variables that relate mainly to the characteristics of borrowers. While the focus of the study is on the borrower's attributes, LTV ratio and loan duration (DURA) have been included in the model due to their potential influence on loan default. The variable LTV measures the relationship between the loan and the value of the real property which is the main form of collateral for the loans under study. Based on extant studies, it is hypothesized that high LTV would result to high probability of default. The variable DURA is included as a potential determinant of loan recovery on the expectation that long-term loans would have high probability of default than short-term loans. The variable PTI is the ratio of a borrower's loan repayment to income. It is hypothesized that the high PTI ratio would result to high probability of default and vice versa.

Type of borrower (BTYPE) is used to distinguish those who are taking loan for the first time and those who have taken loans before regardless of the purpose. The variable tests if being a first-time borrower has any impact on probability of default. It is hypothesized that borrowers who have engaged in lending in the past would have good experience with respect to loan repayment which tends to help them not to default. The variable HIST is used to define whether a borrower's credit history is available to the lender or not as at the time of loan application. The variable tests if the ability of a borrower to produce evidences on past loan repayment has significant effect on probability of default. Such evidences may include but not limited to attestation or testimonies from former lender(s) on the credibility of the borrower, disposition to loan contract terms and ability to repay. The authors acknowledge that while a borrower's present or future action may not necessarily be determined by past behaviour, it is reasonable to suggest that borrowers who are bold enough to produce a (good) credit history tend to be more trustworthy than those who are not able. It should be stated though that credit history does not apply to first-time borrowers. So, a similar variable coded as GUARAN has been introduced. Similar to HIST, the variable tests if availability of a credible guarantor has significant influence on the probability of default.

Furthermore, the variable EMP is included in the model to test if borrower's employment status has significant impact on odds of default. The variable has three categories (EMP1 to EMP3) each representing the employment status of borrowers. It is hypothesized that probability of default would be lower among the salary earners compared to the self-employed since PTI ratio can be adequately determined at loan origination. In addition, income volatility tends to be higher among the self-employed than the salary earners which may further increase the odds of default. Also, the variable BSEX is included in the model to test if the gender category of a borrower has significant impact on probability of default. The borrower's marital status (MSTAT) divided into three categories (MSTAT1 to MSTAT3) captures the impact of borrower's marital status on probability of default. Literature is not clear on whether a borrower's marital status has significant effect on probability of default. However, it is hypothesized in this study that there will be difference in the performance of the three marital statuses in respect of mortgage obligation with the married expected to have the least probability of default.

The variable AGE represents the age of borrower although the impact of age on loan default may be difficult to predict. For instance, on the one hand, older borrowers are expected to be responsible and mature, and so should not be prone to default. However, they tend to also have more dependants and similar responsibilities that may impact on their mortgage commitment. The relationship between age and loan default may not therefore be linear. The borrower's highest educational attainment (EDU) is measured as a dummy variable with higher education degree taken as the reference category. Based on the assumption that education and level of income are positively correlated, it is hypothesized that probability of default should be lower among borrowers who are graduates than their counterparts with lower educational attainment. Another factor that may have significant effect on a borrower's probability of default is number of dependants. The variable DEPEND is therefore included in the model to test if the number of dependants as at date of loan origination is a significant factor of default.

The last variable used for the model development is the borrower's tribe (TRIBE) which has three categories (TRIBE1 to TRIBE3) each representing the three main tribes in Nigeria. Similar to studies that have found race as a determinant of mortgage default, TRIBE is included in the model to capture the effect of tribe on default. In this study however, the spatial implication of tribe is the main focus rather than an assessment of commitment to mortgage obligation across the tribal groups. Thus, tribe is taken as a measure of distance from place of origin and perhaps extended families and relations. It is assumed that borrowers who live farther away from their tribal region and place of origin are likely to be less prone to mortgage default than their counterparts who live and settle among their acquaintances and tribal people. Since Lagos metropolis is the study base and falls within the Yoruba tribal region, it is hypothesized that the Yorubas should be more prone to loan default than their Igbo and Hausa counterparts.

Using an LR model as specified in Equation (1), loan default is expressed as a function of financial and non-financial characteristics of borrowers. The central mathematical concept that underlies LR is the logit – the natural logarithm of an odds ratio (Peng, Lee, & Ingersoll, 2002).

$$\text{Logit}\,(\text{LNDFT} = 1|X_1,\ldots\ldots,X_n) = \ln\left(\frac{P_i}{1-P_i}\right) = \alpha + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n \quad (1)$$

where P_i is the probability of defaulting denoted as LNDFT = 1, X_1, ……, X_n are the independent variables which are mainly borrowers' attributes, α is the intercept, β_1, ……, β_n is the regression coefficient.

In order to derive the equation to predict the probability of the occurrence of the outcome of interest, the antilog of both sides of Equation (1) is calculated as in Equation (2).

$$P_i = \text{Probability}\,(\text{LNDFT} = 1|X_1,\ldots\ldots,X_n) = \frac{e^{\alpha+\beta_1X_1+\beta_2X_2+\ldots+\beta_nX_n}}{1+e^{\alpha+\beta_1X_1+\beta_2X_2+\ldots+\beta_nX_n}} \quad (2)$$

It should be stated that the software (SPSS) used for the analysis generates the antilog of the coefficients which is the effect of a unit change in the value of an independent variable on the odds of default. In order to estimate the effect of each independent variable on the odds of default, 1 is subtracted from odds and the difference is multiplied by 100.

Results and discussion

Data were analysed using frequency distribution (Table 2), independent *t*-test (Table 3), and LR (Table 5). Frequency distribution and descriptive statistics are used to present the characteristics of the borrowers/loans under study. Independent *t*-test is then used to assess the difference between the populations of defaulting and non-defaulting

Table 2. Frequency distributions.

Category	Frequency		Percentage
Loan default status	Defaulted	163	53.40
	Non-defaulted	142	46.60
	Total	305	100.00
Type of borrower	First-time borrower	106	34.80
	Old borrower	199	65.20
	Total	305	100.00
Credit history	Available	146	47.90
	Unavailable	159	52.10
	Total	305	100.00
Employment status	Salary Earners	130	42.60
	Self-employed	175	57.40
	Retired/Pensioner	0	0.00
	Total	305	100.00
Borrower's sex	Male	235	77.00
	Female	70	23.00
	Total	305	100.00
Marital status	Married	295	96.70
	Single	9	3.00
	Widowed/Divorced	1	0.30
	Total	305	100.00
Educational status	Graduate	227	74.40
	Non-graduate	78	25.60
	Total	305	100.00
Availability of guarantor	Available	269	88.20
	Unavailable	36	11.80
	Total	305	100.00
Borrower's tribe	Yoruba	177	58.00
	Igbo	96	31.50
	Hausa	32	10.50
	Total	305	100.00

borrowers. Lastly, LR is developed to predict mortgage default based on the characteristics of borrowers.

Table 2 shows that 163 borrowers out of the total population of 305 are defaulting while the remaining 142 borrowers are not. With almost equal proportion, the bias that may be introduced as a result of too much gap between the two groups is highly reduced. Among the borrowers, 106 are obtaining loan for the first time, while 199 borrowers had obtained loan before. The table also shows that the credit history of 146 borrowers against 159 was available to the lenders at loan application. The employment status is divided into three categories and as shown in the table, there are no retired/pensioners among the borrowers. More than half (57%) of the borrowers are self-employed, while 130 borrowers (representing 43%) are salary earners. There are more male (235) than female (70) among the borrowers. Also, majority of the borrowers (97.6%) are married, while the singles and the widowed/divorced are in the minority representing 3 and 0.3%, respectively. There are 227 graduates (representing 74%) among the sampled borrowers, while 78 borrowers (representing 26%) are not educated up to a first-degree level. Lastly, the table shows that majority (58%) of the borrowers are Yorubas, while both the Igbos (32%) and the Hausas (11%) are in the minority.

Table 3. Comparing the characteristics of the defaulting and the non-defaulting mortgagors.

	Non-defaulting ($N = 142$)		Defaulting ($N = 163$)		Independent t-test statistics		
Variable	Mean	Std. Dev.	Mean	Std. Dev.	Mean difference	t-value	Sig. (2 tailed)
LTV[1]	0.55	0.18	0.54	0.15	−0.01	−0.481	0.631
DURA[2]	29.71	12.58	28.62	10.79	−1.09	−0.807	0.420
PTI[1]	0.55	0.19	0.61	0.20	0.07	2.998	0.003**
BTYPE[2]	0.22	0.42	0.46	0.50	0.24	4.616	0.000**
HIST[1]	0.53	0.50	0.44	0.50	−0.09	−1.616	0.107
EMP1[1]	0.38	0.49	0.47	0.50	0.09	1.515	0.131
EMP2[2]	0.61	0.49	0.54	0.50	−0.07	−1.284	0.200
EMP3	0.00	0.00	0.00	0.00	0.00	[a]	[a]
BSEX[2]	0.73	0.45	0.81	0.39	0.08	1.737	0.083
MSTAT1[2]	0.99	0.12	0.95	0.22	−0.04	−1.780	0.076
MSTAT2[2]	0.01	0.08	0.05	0.22	0.04	2.288	0.023*
MSTAT3[2]	0.01	0.08	0.00	0.00	−0.01	−1.000	0.319
AGE[1]	43.57	6.56	43.29	6.40	−0.28	−0.380	0.704
EDU[2]	0.77	0.42	0.72	0.45	−0.06	−1.140	0.255
DEPEND[2]	3.67	1.25	3.40	1.42	−0.26	−1.725	0.086
GUARAN[2]	0.94	0.25	0.83	0.37	−0.10	−2.865	0.004**
TRIBE1[2]	0.54	0.50	0.62	0.49	0.08	1.488	0.138
TRIBE2[1]	0.33	0.47	0.30	0.46	−0.03	−0.568	0.570
TRIBE3[2]	0.13	0.34	0.08	0.27	−0.05	−1.514	0.131

[a] t is not computed because the standard deviations of both groups are zero.
**Significant at 0.01 level;
*Significant at 0.05 level.
N = total population in each group.
[1]Variable = t-test based on "equal variances assumed" as indicated in the Levene's test when F-value is not significant at 0.05 level.
[2]Variable = t-test based on "equal variances not assumed" as indicated in the Levene's test when F-value is significant at 0.05 level.

Table 3 presents the result of the unpaired t-test to analyse the distinguishing factors between the defaulting and the non-defaulting borrowers. The t-tests indicate that there is difference at various significant levels, between the two categories in only four variables. The two categories differ in PTI ratio with defaulting group having a significantly higher value (0.61) than their non-defaulting counterparts with 0.55 (at $p \leq 0.01$). The implication of this is, the higher the PTI ratio, the higher the possibility of default. The table also reveals a significantly higher number of first-time borrowers among the defaulting group than among the non-defaulting (at $p \leq 0.01$). This implies a higher possibility of default among borrowers who have never taken loan before. Also, the number of borrowers whose loans are backed by guarantors is significantly higher among the non-defaulting borrowers than among the defaulters (at $p \leq 0.01$). It can be observed that there are more singles among the defaulting borrowers than among the non-defaulters. As shown in the table, other variables are not differentiating factors between the two borrower categories except EMP, BSEX, DEPEND and TRIBE if the level of significance is ($p \leq 0.1$).

Descriptively, the average LTV of the sampled loans is 0.54 which implies that loan size is about a half of the value of the collateralized real property. The table also shows that the average loan duration is 29 years, while in terms of age, the sampled borrowers are approximately 40 years old with an average of four dependants. The mean value of 0.58 for PTI ratio implies that the borrowers spend on the average 58% of their gross monthly income in repaying their debt. As observed from the data analysis, it appears that the rules guiding mortgage application are not strictly followed with reference to purpose of loan. For instance, some loans are found with duration as low as 10 years which is unrealistic for mortgages. Although the number is not huge enough to affect the result of the analysis, it points to the fact that the lending institutions may be engaging in other activities other than mortgage lending. Omirin and Nubi (2007) note that for pecuniary reasons, several PMIs engage in direct construction of houses for sale thereby competing with the operators they are expected to be financing. In addition, they also engage in non-housing businesses, while some of them are accustomed to giving out short-term loans to traders (Omirin & Nubi, 2007).

Prior to the LR analysis, various tests were applied to evaluate the model in order to confirm if the model fits well as a loan default prediction tool. These are presented in Table 4. The Omnibus test of model coefficients is used to check that the new model (with all explanatory variables included) is an improvement over the baseline model. The test is based on χ^2 and is aimed at determining if there is a significant difference between the log-likelihoods (−2LLs) of the baseline model and the new model. If the new model has a significantly reduced −2LL compared to the baseline, then it suggests that the new model is explaining more of the variance in the outcome and is an improvement. As shown in the table, the chi-square is highly significant ($\chi^2 = 47.69$, $df = 16$, $p < 0.000$) which implies that the new model performs significantly better.

The model summary presents the −2LL and pseudo R^2 values for the model (Table 4). The −2LL value for this model (373.68) is what was compared to the −2LL for the previous null model in the Omnibus test which shows that the new model is significantly better fit than the baseline model. The R^2 values reveal approximately how much variation in the outcome is explained by the model but not exactly the way R^2 in linear regression analysis works. The R^2 values in the context of LR are mere approximations and should not be overly emphasized. Lastly, the Hosmer and Lemeshow test of the goodness of fit suggests that the model is a good fit to the data as $p = 0.872$ (>0.05). However, it should be stated that the χ^2 statistic on which the test is based is very dependent on sample size. It therefore means that its value should not be interpreted without due consideration for the size of the sample used for the analysis.

Table 5 presents the empirical result of the LR which evaluates the variables that influence probability of default. As shown in the table, the only influence variable among the independent variables (at $p \leq 0.01$) is BTYPE, while at $p \leq 0.05$, the influence variables are PTI and BSEX. Other variables show weak or insignificant influence on the dependent variable. The odds ratios of the variables are contained in

Table 4. Evaluations of the logistic regression model.

Test				
Omnibus tests		χ^2	df	Sig.
	Step	47.69	16	0.000
	Block	47.69	16	0.000
	Model	47.69	16	0.000
Model summary		-2LL	Cox & Snell R^2	Nagelkerke R^2
		373.68	0.15	0.193
Hosmer and Lemeshow test		χ^2	df	Sig.
		3.830	8	0.872

the column “Exp (B)” which is the antilog of the coefficients. It represents the effect of a unit change in the value of an independent variable on the odds of default controlling for all other variables. In order to estimate the effect of each of the independent variables on odds of default as in the column “Effect (%)”, 1 is subtracted from odds and the difference is multiplied by 100. It should be stated that in LR, the operational sign (−/+) on the coefficients of independent variables is of great importance because it goes a long way in determining the nature of relationship between dependent variable and the independent variable. While inverse (negative) relationship is depicted by minus, plus depicts direct (positive) relationship. For instance, Table 6 shows that LTV has negative sign which implies that log of default increases with decrease in the value of the former. This also means that probability of default is high when loan size is comparatively larger than the value of collateral. Even though the relationship is not statistically significant, the result is meaningful in mortgage default analysis when the operational sign on the coefficient is considered.

The prediction performance of the LR model is presented in Table 6. The classification table measures the accuracy of the developed model as a tool for predicting loan default. As shown in the table, the model is capable of achieving 68.2% classification accuracy. When this is compared to the equivalent statistics (53.4%) in the null model, a significant improvement is achieved by the actual (developed) model. This gives an indication of the significant explanatory power of the variables included in the final model. Notwithstanding, as this alone may not be good enough to measure accuracy, the overall percentage of the developed model has been compared to past studies as documented in Abdou and Pointon (2011). The comparative assessment revealed that the prediction accuracy achieved by this model ranks well compared to similar studies.

Table 5. Mortgage default factors (Logistic regression result).

Variables	*B*	Wald	Sig.	Exp (B)	Effect (%)
LTV	−0.595	0.598	0.439	0.551	−44.90
DURA	0.010	0.764	0.382	1.010	1.00
PTI	1.484	4.756	0.029*	4.412	341.20
BTYPE	1.148	10.624	0.001**	3.151	215.10
HIST	0.381	1.585	0.208	1.464	46.40
EMP1	21.764	0.000	0.999	2.83 E + 09	2.83 E + 11
EMP2	21.812	0.000	0.999	2.97 E + 09	2.97 E + 11
EMP3	Dropped				
BSEX	0.785	5.659	0.017*	2.193	119.30
MSTAT1	20.401	0.000	1.000	7.25E + 08	7.25E + 11
MSTAT2	21.930	0.000	1.000	3.34 E + 09	3.34 E + 11
MSTAT3	Dropped				
AGE	−0.006	0.078	0.780	0.994	−0.60
EDU	−0.458	2.392	0.122	0.633	−36.70
DEPEND	−0.157	2.069	0.150	0.855	−14.50
GUARAN	−0.548	1.260	0.262	0.578	−42.20
TRIBE1	0.610	1.915	0.166	1.840	84.00
TRIBE2	0.668	2.093	0.148	1.949	94.90
TRIBE3	Dropped				
Intercept	−43.001	0.000	0.999	0.000	NA

**Significant at 0.01 level; *Significant at 0.05 level.
NA = Not applicable.

Table 6. Classification table.

Model	Observed		Predicted: Default status: Non-default	Default	Percentage correct
Null model	Default status	Non-default	0	142	0.0
		Default	0	163	100.0
	Overall percentage				53.4
Actual model	Default status	Non-default	97	45	68.3
		Default	52	111	68.1
	Overall percentage				68.2

Discussion of findings

The findings of this study as presented in the preceding section agree with a number of extant literatures. As observed from the results of both the independent *t*-test and LR analyses, PTI ratio is a significant factor of mortgage default among the borrowers investigated. The variable captures the relationship between monthly loan repayment and the income of borrowers. With this finding, it shows that when the monthly repayment is large compared to income, a borrower tend to find it difficult to fulfil the required repayment obligation. As found, increasing PTI by one unit will increase the odds of default by approximately 341%. This finding is in line with the findings of most previous studies reviewed in this study (Canner et al., 1991; Lee & Liu, 2001). While PTI alone may not be responsible for the failure of borrowers to meet repayment obligations, this finding suggests that income plays an important role in determining probability of default as established by Peter and Peter (2006).

Another variable that has significant impact on mortgage default as observed in the results of both inferential statistical analyses is BTYPE. The variable captures difference in the probability of default between those who are first-time loan takers and those who had previously taken loans. As observed from the LR result, first-time borrowers are more likely to default thrice their counterparts who had obtained loan in the past. While much is not said about the potential relationship between default and being new to borrowing, it is rational to suggest that experience plays an important role in loan repayment. This agrees with the opinion of Afolabi (2010) that borrowers who are experienced in what they are using their loans for are unlikely to default. This suggests that borrowers who are obtaining loans for the same lines of development activities tend to accumulate experience that would assist them to be successful and hence be able to repay. In addition, past experiences garnered by the old borrowers can be useful in devising proactive strategies to achieve timely repayment.

The sex category of borrowers is also found to be a significant determinant of mortgage default as indicated in the result of the logistic model. The result shows that the default probability of male borrowers is significantly higher than that of females in accordance with the finding of Lin et al. (2011). As indicated, being a male borrower will increase the odds of default two times more than being a female. Practically, this seems to be inconsistent as males are expected to be generally more educated and financially buoyant than females. This finding suggests that the tendency of females to be naturally tender-hearted may be a motivation for their lower probability of default. However, according to Lin et al. (2011), male borrowers are usually more burdened with family responsibility than females which may have significant consequence on their ability to fulfil mortgage obligations.

Based on the result of the independent *t*-test, availability of guarantor (GUARAN) and whether a borrower is single or not (MSTAT2) are also found as potential variables of difference between the defaulting and the non-defaulting borrowers. A higher mean value of GUARAN (0.94) for non-defaulters against 0.83 for the defaulting borrowers gives an indication that the number of borrowers whose loans are backed by guarantors is significantly higher among the non-defaulting borrowers than among the defaulters. Although this variable is not significant in the LR result, probability of default tends to reduce when borrowers are made to provide guarantors. It is generally the case that borrowers tend to maintain continuous cordial relationships with guarantors by not defaulting. On the other hand, repayment responsibility can also be shifted to guarantors in case of default.

Furthermore, based on the same *t*-test result, it is observed that there are more unmarried among the defaulting borrowers than among the non-defaulters. This has certain degree of similarity with past research findings (Lin et al., 2011; Peter & Peter, 2006). Although the variable (MSTAT2) is not significant in the LR, it suggests that being single is related to high default probability. Practically, the unmarried tend to be young in terms of age and natural risk-takers. Age is generally accepted as a measure of maturity and experience which means that the higher number of singles among the defaulting borrowers can be as a result of their low experience in life. On the other hand, singles as natural risk-takers tend to engage in risky investments that promise bumper return. Sometimes, this could even be in form of diverting the loan to a venture that is completely different from its original purpose. Unfortunately for these borrowers, such investment decisions do not always turn out positive and the consequence is usually failure to meet repayment obligations. Therefore, these two characteristics can help explain the high propensity of singles to default on mortgages.

Nonetheless, there are certain degree of dissimilarities between the current study and the reviewed empirical researches. For instance, educational status, age and employment are found as significant factors of influence on mortgage default probability (Bandyopadhyay & Saha, 2009; Lin et al., 2011; Peter & Peter, 2006). However, in this study, these variables seem not to have significant impact on mortgage default, although the direction of their influence on probability of default is similar to previous studies. For instance, highest educational attainment (EDU) is found to be an insignificant factor of default. It should be recalled that the variable is adopted as one of the borrowers' income parameters based on the assumption that being a graduate should mean more money compared to other lower educational attainment. With this finding, however, income seems not to be a function of education, hence its little explanatory power in the LR equation.

The finding on EDU is also applicable to the insignificant explanatory power of employment (EMP) categories on the probability of default. It implies that being a salary earner, self-employed or retired does not have any impact of a borrower's ability to repay. As observed in the results too, increasing AGE by one unit reduces the probability of default by 0.6%. This seems to disagree with the explanation made on the likely relationship between age and being single which means that further investigation is needed to uncover how these variables contemporaneously impact on mortgage default. Lastly, while Afolabi (2010) found that family size and expenses have negative impact on loan repayment, this study seems not to support this as the average number of dependants in non-defaulting group is similar to that of the defaulting category.

Conclusion

The need for proper identification and evaluation of borrower-specific characteristics as loan default triggers has been stressed in this study. This is in response to the debilitating impact of default on lending market and real estate finance in particular as underscored in recent studies. No doubt, the financial capabilities and general creditworthiness of borrowers are key factors in assessing default risk (Hayre, Saraf, Young, & Chen, 2008). Hence, developing a model to evaluate the potential impact of borrower's characteristics on loan default is a step in the right direction towards effective credit risk evaluation. Building on existing studies, the current study assessed the impact of borrower's attributes on default probability. The study applied LR to evaluate the general characteristics of borrowers as possible explanatory variables for default in residential mortgages. Borrowers' socio-economic attributes, namely PTI ratio, BTYPE and sex are found as significant factors inducing default among the sampled borrowers. Based on *t*-test results, whether a borrower is single or not and whether a loan is backed by guarantor or not are also found as key factors of default. While the findings of this study share some commonalities with extant literatures, a number of inconsistencies are noticed which call for further research.

For instance, it is not clear how age and marital status relate to determining a borrower's default probability. More evidence is also needed on the relationship between education and income and how they jointly act to influence repayment ability. Based on these highlighted ambiguities, the findings of this study signal a complexity that is inherent in employing socio-economic factors for mortgage default model development. This is because they are almost unpredictable and subject to a lot of behavioural manipulations and external intervening factors. Notwithstanding, it does not undermine the power of LR as a loan default probability model. In dealing with the aforementioned challenge, it is needful to stress that a model such as this should be based on a larger pool of data to reduce the magnitude of error. It is envisaged that increasing the sample size can help achieve increased prediction accuracy. Howbeit, access to a large data-set (if existing) is still difficult in most developing countries. This further strengthens the clamour for the development of national database for the management of borrowers' lending records in a developing country such as Nigeria. Access to such data is no doubt useful for further researches in the area of mortgage lending.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

M.O. Bello http://orcid.org/0000-0002-6294-5195

References

Abdou, H. A., & Pointon, J. (2011). Credit scoring, statistical techniques and evaluation criteria: A review of the literature. *Intelligent Systems in Accounting, Finance and Management, 18*, 59–88. Retrieved July 23, 2015, from http://doi.wiley.com/10.1002/isaf.325

Afolabi, J. A. (2010). Analysis of loan repayment among small scale farmers in Oyo State, Nigeria. *Journal of Social Science, 22*, 115–119.

Ajayi, C. A. (1992). An analysis of default factors in residential mortgages of the federal mortgage bank of Nigeria. *African Review of Money and Banking, 2*, 203–215.

Akinlo, O., & Emmanuel, M. (2014). Determinants of non-performing loans in Nigeria. *Accounting & Taxation, 6*, 21–28.

Aluko, B. T. (2007). Implications of the current trend in mortgage valuation practice in Nigeria. *International Journal of Strategic Property Management, 11*, 17–31.

Anderson, R., & VanderHoff, J. (1999). Mortgage default rates and borrower race. *Journal of Real Estate Research, 18*, 279–289.

Bandyopadhyay, A., & Saha, A. (2009). Factors driving demand and default risk in residential housing loans: Indian evidence. *MPRA Paper,* 14352. Retrieved June 15, 2015, from https://mpra.ub.uni-muenchen.de/14352/

Cairney, J., & Boyle, M. H. (2004). Home ownership, mortgages and psychological distress. *Housing Studies, 19*, 161–174.

Canner, G. B., Gabriel, S. A., & Woolley, J. M. (1991). Race, default risk and mortgage lending: A study of the FHA and conventional loan markets. *Southern Economic Journal, 58*, 249–262.

Clauretie, T. M., & Sirmans, G. S. (2003). *Real estate finance: Theory and practice*. Mason, OH: Thomas Learning.

Clayton, J. (2007). Commercial real estate derivatives: The developing US market. *Journal of Economic Perspectives, 32*, 33–40.

Delgadillo, L. M., & Gallagher, A. (2003). Borrowers and mortgage related factors associated with FHA foreclosures. *Family and Consumer Sciences Research Journal, 34*, 204–222.

Diaz-Serrano, L. (2005). Income Volatility and Residential Mortgage Delinquency across the EU. *Journal of Housing Economics 14*, 153-177.

Hakim, S. R., & Haddad, M. (1999). Borrower attributes and the risk of default of conventional mortgages. *AEJ, 27*, 210–220.

Hayre, L. S., Saraf, M., Young, R., & Chen, J. D. (2008). Modeling of mortgage defaults. *The Journal of Fixed Income, 17*, 6–30.

Jacobson, T., & Roszbach, K. (2003). Bank lending policy, credit scoring and value-at-risk. *Journal of Banking & Finance, 27*, 615–633.

Kama, U., Yakubu, J., Bewaji, P., Adigun, M. A., Adegbe, O., & Elisha, J. D. (2013). *Mortgage financing in Nigeria* Central Bank of Nigeria Occasional Paper No. 50 Retrieved June 18, 2015, from http://www.cenbank.org/Out/2014/RSD/Occasional%20Paper%20No.%2050mortagage%20financing%20in%20NIgeria%20Combined.pdf.

Kung, J. Y., Wu, S. Y., Hsu, S. S., Lee, W., & Yang, C. W. (2010). Application of logistic regression analysis of home mortgage loan prepayment and default risk. *ICIC Express Letters, 4*, 325–331.

Lee, S. P., & Liu, D. Y. (2001). An analysis of default risk on residential mortgage loans. *International Journal of Management, 18*, 421–431.

Leishman, C., & Watkins, C. (2004). The decision-making behaviour of office occupiers. *Journal of Property Investment & Finance, 22*, 307–319.

Lin, T. T., Lee, C., & Chen, C. (2011). Impacts of the borrower's attributes, loan contract contents, and collateral characteristics on mortgage loan default. *The Service Industries Journal, 31*, 1385–1404.

Munnell, A. H., Geoffrey, M. B. Tootell., Lynn, E. B. & McEneaney, J. (1996). Mortgage lending in Boston: Interpreting HMDA data. *American Economic Review, 86*, 25–53.

Nubi, T. O. (2000). *Housing finance in Nigeria – Need for reengineering*. Retrieved November 13, 2010, from http://www.housingfinance.org/pdfstorage/Africa

Ojo, A. O. (2009). *The role of secondary mortgage facility in expanding the availability of funds for mortgage finance in Nigeria*. (An MSc Thesis Submitted to the Department of Real Estate and Construction Management). KTH Architecture and the Built Environment, Stockholm.

Okpara, G. C. (2009). A synthesis of the critical factors affecting performance of the Nigerian banking system. *European Journal of Economics, Finance and Administrative Sciences, 17*, 34–44.

Omirin, M. M., & Nubi, T. G. (2007). The role of primary mortgage institutions in housing delivery. *Housing Finance International, 22*, 52–56.

Onyike, J. A. (2007). An assessment of the affordability of housing by public servants in Owerri, Nigeria. *Journal of Land Use and Development Studies, 3*, 21–34.

Peng, C. J., Lee, K. L., & Ingersoll, G. M. (2002). An introduction to logistic regression analysis and reporting. *The Journal of Educational Research, 96*, 3–14.

Peter, V. & Peter, R. (2006). Risk management model: an empirical assessment of the risk of default. *International Research Journal of Finance and Economics, 1*, 42–56.
Saka, N., & Akin, O. T. (2007). An assessment of the impact of real estate sector on the Nigerian economy. *Journal of Land Use and Development Studies, 3*, 155–171.
Sanusi, J. O. (2003). *Mortgage financing in Nigeria: Issues and challenges*. Being Text of a Paper Presented at the 9th John Wood Ekpenyong Memorial Lecture, Organized by the Nigerian Institution of Estate Surveyors and Valuers, 29th January. Retrieved September 9, 2014, from http://www.cenbank.org/out/Speeches/2003/Govsp-29Jan.pdf.
Somoye, R. O. C. (2010). The variation of risks on non-performing loans on bank performances in Nigeria. *Indian Journal of Economics and Business, 9*, 87–99.