$x_1, x_2, x_3, \ldots\ldots x_k$ are a vector of observed covariates (independent variables) and $\alpha, \beta_1, \beta_2, \beta_3, \ldots\ldots, \beta_k$ are a vector of regression coefficients of the independent variables to be determined.

2.2.4 The Logistic Regression Model

Consider k independent observations y_1, y_2, ………………y_k and where the i-th observation is a realization of a random variable Y_i. Assuming $Y_i \sim B(1, \pi_i)$ the logit of the probability π_i is the linear function of

$$\mathrm{logit}(\pi_i) = X'\beta \quad \text{.............................(3.7)}$$

where X_i are a vector of covariates and β_i are a vector of regression coefficients.

From equation 3.7 the odds for the i[th] unit are given by

$$\frac{\pi_i}{1-\pi_i} = \exp(X'_i\beta) \quad \text{........................(3.8)}$$

Solving for π_i in equation 3.8 gives

$$\pi_i = \frac{\exp(x'_i\beta)}{1+\exp(x'_i\beta)} \quad \text{........................(3.9)}$$

This can be re-written as

$$f(y) = \frac{e^z}{1+e^z} \quad \text{..............................(3.10)}$$

Where z is the logit of y defined as

$$z = \alpha + \beta_1 x_1 + \beta_2 x_2 + \cdots \ldots\ldots + \beta_k x_k \quad \text{....(3.11)}$$

The socio-demographic variables of interest in our study were gender, age, marital status and education level of a cardholder.

Incorporating these variables in the logistic regression model defined above gave the general model for the study as

$$\pi_i = \text{f(ag, ms, gd, ed) ....(3.12)}$$

Where

π_i = probability of default in credit card by the i-th cardholder

ag = socio-demographic factors, age

ms = socio-demographic factors, marital status

gd = socio-demographic factors, gender

ed = socio-demographic factors, education level

From equations 3.11 and 3.12

$$z = \alpha + \beta_1 ag + \beta_2 ms + \beta_3 gd + \beta_4 ed \quad \text{...............(3.13)}$$

2.5 Data Analysis

Both descriptive and inferential data analysis were carried out. Chi-square testing for independence of variables was carried out to identify if there were statistically significant associations between categorical variables (gender, marital status and education level) and default in credit cards. For the continuous variable ,age, independent samples t-tests were carried out to obtain the significance in the difference of means for the defaulted and non-defaulted groups under statistical investigation. To draw inferences about the influence on credit cards default by each variable of interest, a logistic regression model was fitted and run in SPSS 20. Marginal effects analysis for the effect of a unit change in the independent variable on credit card default was carried out using the odds ratio.

3. Findings

3.1 Influence of Gender on credit card default

From the results female cardholders had a lower default rate of 13.7% compared with male cardholders whose default rate was 27.4%. Also from the study results, 64.2% of sampled credit cardholders were male. These results are consistent with findings by Abdul-Muhmin and Umar (2007) that the tendency to revolve in credit cards is higher among males.

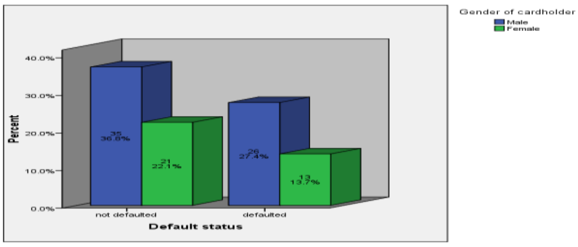

Figure 3.1: Default status by gender of credit cardholder

Despite the observed relatively higher default rate among male cardholders, the Chi-square results showed that there was no statistically significant relationship between gender and credit card default rate ($\chi 2 = 0.174$, $p = 0.677$, $\alpha = 0.05$) which implied that gender, taken alone did not influence default in credit card. These results vary with the findings of Arminger et al., (1997), Kocenda and Vojtek, (2009); Dunn and Kim (1999) that gender is a risk factor in loans and that females default less frequently possibly because they are more risk averse.

3.2 Influence of Age on credit card default

From group statistics of age of cardholders, the study results shows that the mean age of cardholders who defaulted was 44.18 years which was lower than that of non-defaulters which was 52.14 years.

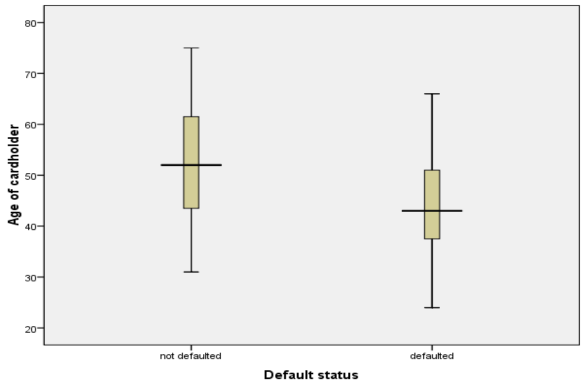

Figure 3.2: Default status by Age of cardholder

The Levene's test of equivalence of variance (p = 0.854, α = 0.05) showed that the variance of the two groups, defaulted and non-defaulted are the same. However, the t-test for equality of means (p = 0.000, α = 0.05) indicated that there was statistically significant relationship at 5% level of significance between age of cardholder and credit card default. In particular, young cardholders had a higher default rate compared to older cardholders. These results are consistent with literature, Dunn and Kim (1999); Arminger et al. (1997) as well as Agarwal et al. (2009) that older borrowers are more risk averse and will therefore be less likely to default.

3.3 Influence of marital status on credit card default

Descriptively the study results show that there was lower default rate of 39.5% among married cardholders compared to cardholders in other marital status whose default rate was 47.4%. However, Chi-square testing shows there was no significant relationship between marital status and credit card default rate ($\chi2$ = 0.391, p = 0.531, α = 0.05) which implied that default in credit card was independent of marital status of the cardholder.

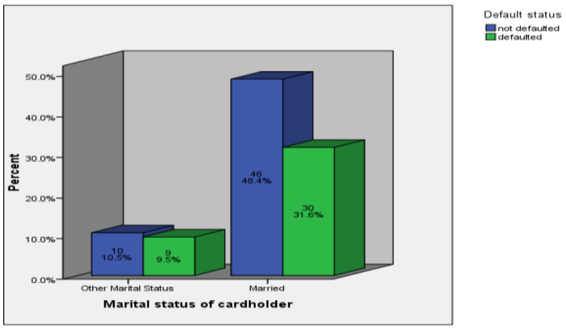

Figure 3.3: Default status by marital status of cardholder

This study result disagrees with the study by Agarwak *et al* (2009) which indicated that marital status can predict default rate on the basis that marital status should be seen to be a sign of responsibility, reliability or maturity of a borrower.

3.4 Influence of Education level on credit card default

On education level, the results showed that 16.8% of cardholders with university level education defaulted compared with 24.2% for cardholders with education level lower than university. These results are presented in Figure 4.4. Consistent with the findings of Steenackers and Goovaerts (1989) that customers who are highly-educated professionals were less likely to default on their credit cards, the current study similarly observed, albeit descriptively, a lower frequency in default for cardholders with university education relative to cardholders with lower than university education.

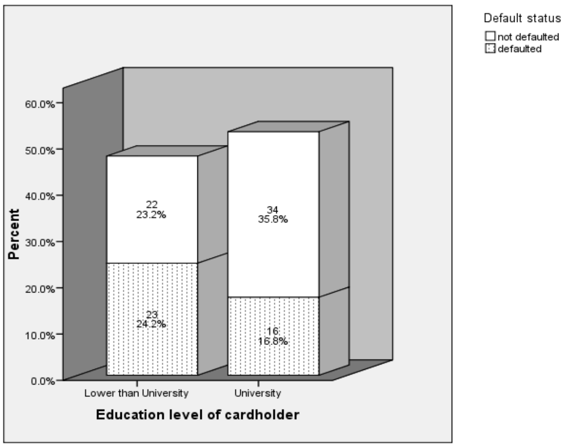

Figure 3.4: Default status by Education level of cardholder

However, Chi-Square tests results ($\chi2$ = 3.575, p = 0.059, α = 0.05) showed that education level of a cardholder is independent of credit card default.

4. Conclusions

A number of socio-demographic factors influence credit cards default. Among them, age has the highest influence and is statistically significant in credit cards default. While gender, marital status and education level also affect credit cards default, their influence is statistically insignificant. The research further shows that each of the socio-demographic factors have strategic significance to credit cards issuers and would be useful in mitigating credit cards default.

5. Future Scope

The current study used logistic regression to investigate the influence of socio-demographic, behavioral and economic determinants on credit cards default in commercial banks in Kenya. The focus of this study was to obtain a set of explanatory variables with highest predictive probabilities of default in credit cards as loan assets. As per the Basel II framework and requirements, future studies may address other components of expected loss for credit cards which includes; Loss Given Default (LGD), Exposure at Default (EAD) and Maturity of Exposures (M). Future studies could also explore use of other statistical techniques such multiple discriminant analysis model, linear probability model or the probit model.

References

[1] Agarwal, S., Chomsisengphet, S. & Liu, C. (2009). Consumer Bankruptcy and Default: The Role of Individual Social Capital. Working Paper. Available at SSRN: http://ssrn.com/abstract=1408757.

[2] Arminger, G., Enache, D. & Bonne, T. (1997). Analyzing Credit Risk Data: A Comparison of Logistic Discrimination, Classification Tree Analysis, and Feedforward Network.*Computational Statistics*, Vol. 12, Issue 2, p. 293-310.

[3] Autio, M., Wilska, T-A., Kaartinen, R. & Lähteenmaa, J. (2009). The Use of Small Instant Loans Among

Young Adults – a Gateway to a Consumer Insolvency. *International Journal ofConsumer Studies*, Vol. 33, Issue 4, p. 407-415.

[4] Bofondi, M. & Lotti, F. (2006). Innovation in the Retail Banking Industry: the Diffusion of Credit Scoring. *Review of Industrial Organization.* Vol. 28, Issue 1, p. 343-358.

[5] Bolton,C.(2009). Logistic regression and its applications in credit scoring. Dissertation.

[6] Boyes, W. J., Hoffman, D. L. & Low, S. A. (2002). An econometric analysis of the bank credit scoring problem. *Journal of Econometrics*, Vol. 40, Issue 1, p. 3-14.

[7] Central Bank of Kenya (2014).Credit Survey Report.Nairobi. Central Bank of Kenya

[8] Crook, J. N., Hamilton, R. & Thomas, L. C. (1983). A Comparison of a Credit Scoring Model with a Credit Performance Model. *The Service Industries Journal*, Volume 12, Issue 4, p.558-579.

[9] Dinh, T. H. T. & Kleimeier, S. (2007). A Credit Scoring Model for Vietnam's Retail Banking Market. *International Review of Financial Analysis*, Vol. 16, Issue 5, p. 571-495.

[10] Dunn, L. F & Kim, T. (1999). An Empirical Investigation of Credit Card Default. Working Paper, Ohio State University, Department of Economics, 99-13.

[11] Gross, D. & Souleles, N. (2001). Liquidity Constraints and Interest Rates Matter for

[12] Consumer Behavior? Evidence from Credit Card Data. *The Quarterly Journal of Economics*, Vol. 117, Issue 1, p. 149-185.

[13] Jacobson, T. & Roszbach. K. (2003). Bank lending policy, credit scoring and value-at-risk. *Journal of Banking & Finance*, Vol. 27, Issue 4, p. 615-633.

[14] Kocenda, E. & Vojtek, M.. 2009. Default Predictors and Credit Scoring Models for Retail Banking. CESifo Working Paper, No. 2862.

[15] Martso,(2010). The determinants of default in consumer market. Dissertation.

[16] Mukaya, C. (2011). The relationship between credit default risk and cardholder characteristics, credit card characteristics, behavioral scoring process among commercial banks in kenya

[17] Musto, D. K. & Souleles, N. S. (2006). A Portfolio View of Consumer Credit. *Journal of Monetary Economics*, Vol. 53, Issue 1, p. 59-84.

[18] Olukunle and Siwangaliso.(20120. Credit usage, Hire Purchase costs, consumer protection in retail institutions in Botswana

[19] Roszbach, K. (2004). Bank Lending Policy, Credit Scoring and the Survival of Loans. *Review of Economics and Statistics*, Vol. 86, Issue 4, p. 946-958.

[20] Steenackers, A. & Goovaerts, M. J. (1989). A Credit Scoring Model for Personal Loans.*Insurance: Mathematics and Economics*, Vol. 8, Issue 1, p. 31-34.

[21] Stiglitz, J. E. & Weiss, A. M. (1981). Credit Rationing in Markets with Imperfect Information. *American Economic Review*, Vol. 71, Issue 3, p. 393-410.

[22] Straka, J. W. (2000). A Shift in the Mortgage Landscape: The 1990s Move to Automated Credit Evaluations. *Journal of Housing Research*, Vol. 11, Issue 2, p. 207-232.

[23] Thomas, L. C. (2000). A Survey of Credit and Behavioural Scoring: Forecasting Financial Risk of Lending to Consumers. *International Journal of Forecasting*, Vol. 16, Issue 2, p. 149-172.

[24] Updegrave, (1987). How lender size you up. *Money*, p. 23-40.

[25] Wafula and Karumba.(2012). Collateral Lending: Are there alternatives for the Kenyan banking industry. Wps/03/12

[26] Özdemir, Ö. & Boran, L. (2004). An Empirical Investigation on Consumer Credit Default Risk. Turkish Economic Association Working Paper 2004 / 20.