Journal of Finance and Economics, 2018, Vol. 6, No. 6, 242-249
Available online at http://pubs.sciepub.com/jfe/6/6/6

DOI:10.12691/jfe-6-6-6

A Hidden Markov Model of Risk Classification among the Low Income Earners

Davis Bundi Ntwiga[1,*], Carolyne Ogutu[1], Michael Kiura Kirumbu[2], Patrick Weke[1]

[1]School of Mathematics, University of Nairobi
[2]School of Sciences and Engineering, Daystar University
*Corresponding author: dbundi@uonbi.ac.ke

Received September 30, 2018; Revised November 20, 2018; Accepted December 18, 2018

Abstract Low income earners have volatile incomes and most financial providers shun this group of borrowers even though they are motivated in managing the limited resources they have through savings and investments as a means to lower the fluctuations of their income. Peer groupings of the low income earners can assist in pooling the resources they have and improve the group risk mitigation process as group members act like social collateral in credit lending. The study used Kenya Kenya Financial Diaries data of 2013 from 280 households to analyze and understand the credit quality levels and credit scores of peer groups versus individuals among men and women. Hidden Markov model classified the low income earners into credit risk profiles wih a view of understanding the role of groups in low income group lending. Peer groups diversify risk inherent in individual borrowers with women only groups having higher credit quality levels as compared to men only groups. Women and their respective peer groups are more stable with less variability as compared to men. Financial technology providers can incorporate the wide array of soft information to lend to low income earners through mobile based peer groups.

Keywords: *credit score, hidden Markov model, men, women, peer groups, credit quality, risk classification and low income earners*

Cite This Article: Davis Bundi Ntwiga, Carolyne Ogutu, Michael Kiura Kirumbu, and Patrick Weke, "A Hidden Markov Model of Risk Classification among the Low Income Earners." *Journal of Finance and Economics*, vol. 6, no. 6 (2018): 242-249. doi: 10.12691/jfe-6-6-6.

1. Introduction

In November 2012, Safaricom Kenya and Commercial Bank of Africa started a mobile based Micro-credit facility dubbed M-shwari based on the M-Pesa platform. M-Pesa has become one of the most successful mobile phone based financial service provider in the developing world [1,2]. The product offers a combination of savings and access to micro-credit loan and the target market was to capture both the banked and unbanked in an effort to increase financial inclusion. This product limit the ability of the low income earners to access the micro-credit facility as they lack good financial options. Their income is volatile, fluctuating daily and they lack reliable ways to harness the power of their low income [3,4]. One solution is to form mobile based peer groups for micro-credit lending.

The World Bank Findex data reveal that in the year 2015, around 55 percent of the global unbanked are women. In the developing countries, 59 percent of men have an account compared with 50 percent of women [5]. This gap can be closed with imaginative service design and delivery that account for the economic and social realities that are common among both men and women but with special emphasis to women, who are more marginalized in formal financial services [5,6,7]. Low income earners are normally more highly motivated in managing their limited resources as they save, invest and even re-invest as a means to lower the fluctuations in their income. Women tend to operate in horizontal social networks that increases their social collateral and this lowers moral hazard in the peer group lending setups they form [5,8]. The value of information of each other in a peer group improves the group risk mitigation and the peer groupings in credit lending acts as social collateral [9].

Low income earners are more inclined to financial service providers with whom they can develop long term relationships. A provider can attract this group of customers with products that offer interdependence, uplifts and values the support system common among the low income people [5]. Loan products are designed in a manner that the lender and borrower are willing to adapt to the unpredictable circumstances in terms of repayment due to the shocks low income earners experience on daily basis. Thus, products for low income earners must resonate with the local systems for resource exchange and upliftment [5].

The success of the mobile financial services, or financial technology (FinTech), low cost of collecting data using mobile technology, a need to further increase financial inclusion, availability of advanced data mining techniques and a wide array of data on consumers can offer an alternative source of financial services to the low income earners in Kenya. Financial inclusion is critical in poverty reduction, increases inclusive economic growth [10] and increases as individuals participate to invest in education, manage risks, absorb financial shocks, increase savings and boost productive investments [10,11]. FinTech geared toward mobile wallets for retail market lending can cover this gap. Benefits of FinTech like M-shwari abounds though some are skeptical due to perceived risk and technological failure [12,13].

In this paper, we propose a credit risk classification model that is tested using historical data among the low income earners. The system is based on hidden Markov model (HMM) technique to estimate the credit risk profiles for individual men and women, and the peer groups formed for saving and borrowing purposes. The credit scores and the credit quality levels form the credit risk profile for risk classification. The low income earners are grouped based on gender as men/women, men only and women only. The objective is to analyze the credit scores of peer groups and individuals with an aim of lending to low income earners using mobile micro-credit based peer groups that are operated and managed by the peer members.

The rest of the paper is organized as follows. Section 2 presents the literature review. Section 3 has the methodology, section 3.1 the characteristics of the hidden Markov model, section 3.2, the credit scoring model and section 3.4 the group borrowers model. In section 4, we have the results and discussions and section 5 the conclusions of the study.

2. Literature Review

The perceptions about mobile banking and technology determine the rate of adoption as income alone is not a sufficient indicator [14]. Around 17.4 percent of the adult population in Kenya are financially excluded, where an individual is more likely to use a mobile phone followed by an informal group, a bank, Savings and credit organization and then a Micro-Finance institution for financial services [6]. A need exists for low income earners to access financial services as studies on micro-credit facilities observe that a positive change in income is observable when the poor use these facilities [15,16]. FinTech can alter the lending landscape as banks penetration in Sub-Saharan Africa is only 35 percent and approximately 80 percent of Africa's one billion population lack access to formal bank services [17]. FinTech advances has provided new possibilities, challenges and opportunities to boost financial, credit access and promote financial depth [18].

The FinAccess Household Survey of 2016 noted that the main reasons people join a group are: access to a lump sum for emergencies, daily needs, social reasons, keep money safe, acquire a lump sum for investment and commitment to save [6]. The best indicators of financial access are the number of people, firms and households receiving credit and using other financial products from the financial service providers [13]. Financial diaries study observed that financial service providers should cater for low income earners by providing services and products that cater to small and inconsistent incomes; offer tools to manage daily risks and transactions; assist women better leverage their social networks; accessibility and support the women to face major life transitions [5].

Women have a higher rate of using informal financial services at 51.4 percent compared to men at 30.9 percent while in using mobile financial services, the difference between men (76 percent) and women (67 percent) is 8 percent [6]. Women's networks tend to be horizontal when compared to men whose social networks are vertical. When compared to men, women's financial aspirations and behaviours are heavily colored by social and cultural norms [5]. The gap of lending to low income earners can be reduced if the social collateral possessed by women can help to reduce the moral hazard [8]. People may have access to financial services but do not use them as such services may be un-affordable, unsuitable for their needs, fear of rejection of the users by the service providers, and likelihood of the service providers unwilling to service that market segment due to their high risk profiles [12,19].

The hidden Markov model (HMM) is the classification method in this study and the technique is increasingly becoming popular due to its strong mathematical structure and theoretical basis for wide applications. There is no theoretical limitations of HMM given enough hidden and observation distribution with sufficient training data [21,22] and [20] expanded on HMM as a generalized mixture model with hidden variables and observed emission. A customer relationship dynamic model to estimate the effects of encounters between customers and the firm is highlighted [23]. A default rate model in a bond portfolio by [24] is based on HMM and [25] used HMM for a credit card transaction processing and this drastically reduced the number of false positive transactions. A HMM model [26] for customer credit risk captured the dynamics and observed the sudden downgrading of the customer. A HMM model forecasted the stock market prices and showed a 100 percent accuracy rate in prediction [22]. A default rate HMM based on social media data had an accuracy rate of between 53 percent and 73 percent [27] and [28] applied a HMM approach to lending to peer groups.

Mobile phone usage data predicted loan repayment in a developing country, where the bank can reduce defaults by 41 percent while still accepting 75 percent of the borrowers [29]. The influence of socio-demographic characteristics on credit risk of consumers indicates that males have higher probability of default compared to female [30]. For the marital status, the married defaults in 24 percent and single customers are likely to default in 36 percent of the cases, which is statistically significant. Younger customers are more likely to default as compared to older ones [30]. Education level significantly influences financial inclusion [6]. Group lending yields higher safeguard to the financial institution as compared to individual lending [31,32]. For the group lending, as the group size increases, efficiency reaches a moderate threshold, where group of four to ten members is optimal for group borrowing [15].

3. Methodology

This section highlights the hidden Markov model classification method, credit scoring analysis, the peer group dynamics and individual household data and the conceptual framework to conceptualize the key variables of the study and how these are interlinked. The data used in this study was extracted from the Kenya Financial Diaries (KFD) project of the year 2012-2013 by the Financial Sector Deepening Kenya (FSDK). The KFD project was designed to improve the understanding of how low income earners in Kenya earn and spend money, which offered new and deeper insights on their financial lives [5]. The study covered 300 households from 14 communities in 5 areas in Kenya: Eldoret, Makueni, Mombasa, Nairobi and Vihiga. It recorded in great detail how they earn and spend their money, share resources and utilize financial devices available in bad and good times.

Sampling during KFD study was purposive and thus limited in making a generalization about Kenya as a whole but offers deep insights on the financial lives of the low income earners. The sample mix was 60-40 rural-urban. The data assists to generate hypotheses difficult to generate from larger quantitative surveys. The diaries questionnaires and interviews were designed to capture information in nine broad domains: well being, household flows, households own production and consumption of food, cash on hand, flows of money, outflows of money, flows of money associated with financial devices, and major events in life. See the Financial diaries study data set, general guidelines and related literature [5,33].

Table 1. Selected variables from the KFD project used in this study

Variables	Frequency of use	Description
Age	Once	Age in years and if of working age
Marital status	Once	Single, Married, Living together, Widowed or separated
Gender	Once	Female or Male
Household size	Once	Number of dependants or house members
Education	Once	Years spent in school
Mobile money flows	Once	Amount received or sent via phone
Mobile money transactions	Once	Number of transactions via phone
Income variability	Once	Changes in income in the time period
Social network	Once	Weak, medium, high
Resources given	Once	Out flows
Resources received	Once	In flows
Income	Monthly	Low, medium, high
Expenditure	Monthly	Low, medium, high
Savings	Monthly	Low, medium, high

Table 1 highlights the variables selected from the financial diaries survey to form the data of this study. The variables indicated with frequency of once are those that are captured once from the financial diaries data set while those with frequency as monthly are obtained through the monthly average value. We selected 280 households for further analysis from the 300 households in the KFD dataset as some data entries were missing or incomplete in the set of the 20 households excluded. The data from the variables in Table 1 was then scaled on $(0,1]$ to convert it into probabilities to ease compatibility with HMM. The 280 households were analyzed as individuals and peer groups based and on gender groupings (men only, women only and men/women). The credit scores and credit quality levels were compared for these different groups and individuals.

The statistical software, Ms Excel, STATA version 14 were used to analyze the data combined with Matlab version 7.1 for developing the HMM analysis toolkit. The socio-economic variables for each of the 280 households are treated as data for individuals or for peer groups. The transition matrix and observation matrix for each individual household and each peer group formed are estimated. The HMM is trained based on the matrices and the HMM output are the credit scores and credit quality levels.

3.1. Hidden Markov Model

We classify the mobile micro-credit customers based on their credit scores and credit quality levels from their socio-economic factors in Table 1. We refer to these factors as the credit scoring factors (CSFs) which are used for HMM learning and training. A transition matrix $A_{2\times2}$ has high and low scores for the customers and the observation matrix $B_{2\times3}$ has the three credit quality levels of low, average and good, the risk classification. The two matrices are derived from the financial diaries data [5] of the customers using the CSFs. A HMM can be characterized by the following: number of states in the model, state transition probabilities, observation probability distribution that characterizes each state, initial state distribution, and the observation symbols [20].

The number of states $(M=2)$ in the model with the set of states denoted as

$$S = \{S_1, S_2\} = \{Low, High\}.$$

The state transition probability distribution The state transition probability distribution $A = \{a_{ij}\}$ where

$$a_{ij} = P[q_{t+1} = S_j \mid q_t = S_i],\ t = 1, 2, ..., T.$$

The number of distinct observation symbols $(K=3)$ per state. We denote the set of observation symbols corresponding to the physical output of the system being modeled as

$$V = \{v_1, v_2, v_3\} = \{L, A, G\} = \{Low,\ Average,\ Good\}.$$

The observation symbol probability matrix $B = \{b_j(k)\}$, where

$$b_j(k) = P[v_k \mid S_j],\ 1 \le j \le M.$$

The initial state probability vector $\pi = \{\pi\}$ where

$$\pi = P[q_1 = S_i],\ 1 \le i \le M.$$

We use the notation $\lambda = (A, B, \pi)$ as the set of parameters of the model in the study for both the individual members and the peer groups formed.

3.2. Credit Scoring and Risk Classification

The probabilistic relationship between the households financial activities, the credit scores and the credit quality level which form the risk classification are estimated by the HMM. The classifications offer insights on who among the low earners, that is, the peer borrowers and individual borrowers have stronger credit scores to qualify for a micro-credit loan. The data from the 280 households using a set of 14 variables (see Table 1) is accumulated together to have enough data for HMM training and learning. The output is the credit scores and credit quality levels of these households.

Let $0 < \beta_i \leq 1$ be the credit score of the i^{th} customer with $i = 1, 2, ..., N$. We express $\beta_i = P(O_1, O_2, ..., O_N \mid \lambda)$ as the credit score from the observations of the $N = 280$ households. Let α_i, $i = 1, 2, ..., n$ be the number of customers in a given credit quality score, where $\alpha = \{L, A, G\}$. The estimation of the credit quality level is based on the credit score which is emitted directly by the HMM. The model emits the credit score and credit quality level at the same time to classify the customers into different risk profiles. Credit scores and credit quality levels are dynamic and changes according to the prevailing conditions. The increase in information about a customer over time can be incorporated in the system to increase the risk classification accuracy of the model. This includes repayment history, saving and withdrawal patterns as well as other data from the mobile phone usage and social media.

3.3. Group Borrowers

Let n_k be the group size of group k with $k \geq 2$. Then, we have the number of customers, $N = n_1 + n_2 + ... + n_k$. Let $0 < \gamma_k \leq 1$ be the credit score of the k^{th} group. The credit score is estimated as, $\gamma_k = P(O_k^1, O_k^2, ...O_k^n \mid \lambda_k)$, where, $\lambda_k = (A_k, B_k, \pi_k)$ estimated from the n_k members in the k^{th} group. Let w_{nk} be the number of peer group members in a given credit quality level, where $w = \{L, A, G\}$. The estimated credit quality level is based on the credit score, the HMM emissions.

The selection of the peer groups was based on both systematic and random sampling using the Matlab computer software. The gender variable is used to separate the sample and offer more insight on the credit scoring for the two groups of households. Descriptive statistics using tables, graphs, frequencies, percentages, standard deviation, coefficient of variation and mean provided more insights on the data. Inferential statistics applied lillietest to test for normality and the data was not normally distributed. Friedman test, a non-parametric, tested for difference between the peer groups using the coefficient of variation as the ranking of data is ordinal. The correlation coefficient measured the linear relationship between the men, women and men/women peer groups credit scores and credit quality levels.

4. Results and Discussions

We compare the individual households and peer groups formed based on gender groupings as men and women, men only and women only. The sample comprises of 57.14 percent men and 42.86 percent women who were the households head. The data was ordered such that the first 160 households were men and the rest, 120 were women.

Table 2 shows the comparison between men and women against their credit quality levels. The percentage of individual women in the average and good credit quality levels is 90.44 percent as compared to 42.5 percent for the men. The grouping of men/women has 57.14 percent in the average and good credit quality levels. The comparison was made based on the proportion of each gender in the sample of study.

In Table 3, the group sizes of the peer groups were varied between 2 to 20 members for men/women groupings, and the credit quality levels or risk classifications are noted in percentages. The percentage of the households in each peer group of the low credit quality level increases from a group size of 2 members to 14 members and a reduction is observed in the average credit quality level. Mixed results are observable in the good credit quality level as the group size increases. The average and good credit quality levels of the groups decreases as group size increases which indicates that there is an optimal group size.

Table 2. Credit quality levels compared against the gender

Credit Quality	Number	Percentage of men/women	Number of Men	Number of Women	Percentage of Men	Percentage of Women
Low	120	42.86	92	28	57.50	9.56
Average	37	13.21	27	10	16.87	59.24
Good	123	43.93	41	82	25.63	31.20
Total	280	100	160	120	100	100

Table 3. Systematic sampling of peer group sizes and the percentage of the groups in the respective risk classification

	Peer Group Size							
Credit Quality	2	4	5	7	8	10	14	20
Low	58.57	61.43	65.00	67.44	68.57	75.00	75.00	71.42
Average	23.57	17.14	13.30	11.62	11.43	10.71	10.00	14.29
Good	17.86	21.43	21.70	20.94	20.00	14.29	15.00	14.29

Table 4. Peer group sizes and percentage of credit quality levels based on gender

	Peer Group Size								
	2			4			8		
Gender	L	A	G	L	A	G	L	A	G
Men	100.00	0.00	0.00	83.33	0.00	16.67	56.25	43.75	0.00
Women	0.00	50.00	50.00	12.50	56.25	31.25	8.33	66.67	25.00

Table 5. Randomly selected peer groups of varying sizes comprising men/women, men only and women only

	Men/Women				Men Only				Women Only			
Group	Credit Quality				Credit Quality				Credit Quality			
Size	f	L	A	G	f	L	A	G	F	L	A	G
2	6	66.67	33.33	0.00	2	100.00	0.00	0.00	6	0.00	50.00	50.00
3	10	10.00	50.00	40.00	13	53.85	46.15	0.00	15	13.33	40.00	46.67
4	10	40.00	50.00	10.00	6	83.33	0.00	16.67	16	12.50	56.25	31.25
5	13	23.08	46.15	30.77	11	63.64	36.36	0.00	12	8.33	58.33	33.33
6	7	0.00	71.43	28.57	11	72.73	27.27	0.00	8	12.50	50.00	37.50
7	13	30.77	53.85	15.38	11	36.36	63.64	0.00	7	0.00	42.86	57.14
8	11	27.27	54.55	18.18	16	56.25	43.75	0.00	12	8.33	66.67	25.00
9	7	14.29	85.71	0.00	8	37.50	62.50	0.00	7	14.29	57.14	28.57
10	14	7.14	78.57	14.29	14	64.29	35.71	0.00	6	16.67	50.00	33.33
11	12	25.00	75.00	0.00	14	35.71	64.29	0.00	15	26.67	46.67	26.67
12	13	15.38	69.23	15.38	10	80.00	20.00	0.00	8	0.00	62.50	37.50
Total	116				116				112			
Mean		23.60	60.70	15.70		62.15	36.33	1.52		10.24	52.76	37.00
*f-frequency												
L-Low, A-Average and G-Good												

Table 6. Friedman test for peer groups with three gender groupings

	Men/Women					Men only					Women only				
Source	SS	df	MS	χ^2	$p > \chi^2$	SS	df	MS	χ^2	$p > \chi^2$	SS	df	MS	χ^2	$p > \chi^2$
Columns	0.7	2	0.35	0.79	0.67	8.1	2	4.05	9.53	0.01	17.1	2	8.55	19.36	0.00
Error	25.8	28	0.92			17.4	28	0.62			9.4	28	0.33		
Total	26.5	44				25.5	44				26.5	44			

Table 4 highlights the dynamics observable in terms of gender when grouped in peer groups. For a group size with 2 members, 100 percent of the women peer groups had average or good credit level as compared to none for the men peer groups. For four members, 16.67 percent of men peer groups and 87.50 percent of the women peer groups had average or good credit quality levels. Eight members in a peer group shows that 43.75 percent of men groups and 91.67 percent of women groups had the average or good credit quality levels. When Table 3 and Table 4 are compared, women only groups had superior credit scores and credit quality levels, followed by men/women peer groups and the last are the men only groups.

Table 5 highlights peer groups with 2 to 12 members and the number of times (frequency) each group was randomly selected. Three gender groupings were considered, men/women, men only and women only groups. The credit quality levels were computed for each grouping, to estimate the percentage of low, average and good credit quality levels. For the category of men/women as members of the peer group, the average and good credit quality observed in the peer group with 12 members, 84.62 percent, 3 members, 90.00 percent and 6 members, 100.00 percent. For the peer groups comprising of men only, the average and good credit quality was observed in the groups with 11 members, 64.29 percent and 7 members, 63.64 percent. The peer group made up of women only had an average and good credit quality in the groups of 8 members, 91.67 percent and 3 members, 86.67 percent. On average, the peer groups comprising of men and women had 76.40 percent of groups in the average and good credit quality levels, the groupings of men only with 37.85 percent and women only with 89.76 percent. Evidently, the women only peer groups had superior performance, followed by men and women groups, then men only groups exhibited the lowest credit quality levels.

In Table 6, we are interested in testing if the group size affects the credit quality level. For the peer groups of men/women, $\chi^2 = 0.79$ $(p = 0.67)$ which is not statistically significant, indicating that the peer group size did not influence the credit quality levels of the households. For the peer groups with men only, the results are statistically significant with $\chi^2 = 9.53, (p = 0.01)$

while for the groups with women only, $\chi^2 = 19.36$, $(p < 0.001)$ which is statistically significant. Peer groups with men only or women only credit quality levels are affected by the size of the peer groups. This is not the case for peer groups of men/women. A possibility is that men and women offer different attributes to the group which reduces variability and thus increasing the group positive attributes.

In Figure 1, the credit quality level of the women only groups shows a marked difference with the men only groups and the men/women peer groups. Individual households had a higher range of credit scores while peer groups had a lower range of credit scores. Thus, peer groups offer more stability when compared to individual households in terms of having lower credit risk. The risk inherent in peer groups is lower than that of the individual households.

In Table 7, the individual households have high variability in the credit scores with a coefficient of variation of between 15.51 percent and 27.78 percent as compared to the peer groups with a coefficient of variation ranging between 5.91 percent and 9.97 percent. Peer groups have more stable credit scores in comparison to the individual households. For the credit quality levels, the average and good credit quality probabilities for the individuals ranges between 62.2 percent and 68.8 percent while the peer groups rating ranges between 64.00 percent and 69.3 percent. Peer groups tend to have lower credit risk or higher credit scores as compared to the individual households.

In Table 8, there exists a difference between the individual households and the peer groups. Peer groups and individual households coefficient of variation differences were tested using Friedman test, they had $\chi^2 = 13.1$, $(p = 0.023)$ which is statistically significant. For the individual households and peer groups, the respective $\chi^2 = 0.67, (p = 0.7165)$ and $\chi^2 = 4.67, (p = 0.097)$ are not statistically significant. The linear relationship measured by correlation coefficient between men only groups and women only peer groups had $r = 0.07895$, men only with men/women peer groups, $r = 0.5171$, and women only with men/women peer groups, $r = 0.5137$. The peer groupings had higher influence on the linear relationship between the men only and women only peer groups.

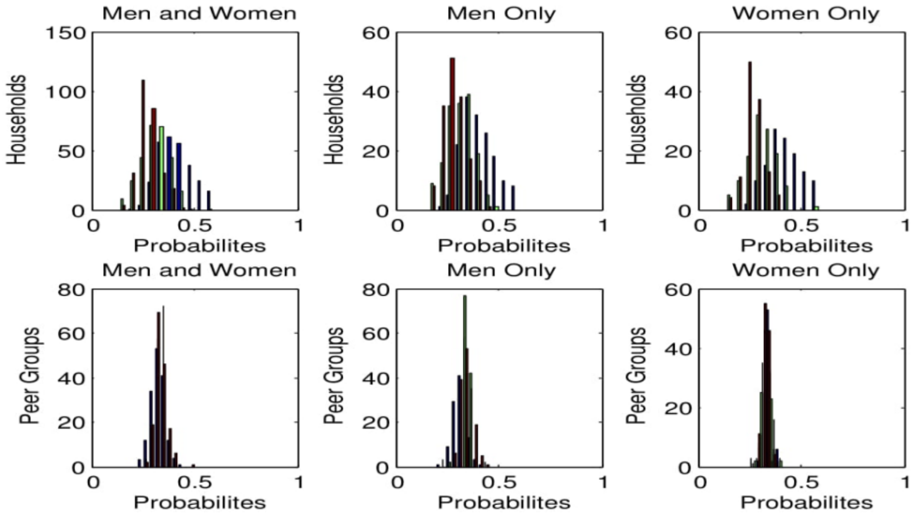

Figure 1. Credit scores of the peer groups and individual households with gender grouping

Table 7. Descriptive statistics for the credit scores and credit quality levels of the peer groups and individuals

Individual Households									
	Men/Women			Men only			Women only		
Statistic	L	A	G	L	A	G	L	A	G
Mean	0.350	0.295	0.355	0.378	0.301	0.321	0.312	0.287	0.401
Stdev	0.088	0.061	0.081	0.078	0.056	0.076	0.087	0.066	0.062
CV (%)	25.10	20.56	22.68	20.49	18.47	27.78	23.08	15.51	18.700
Skewness	-0.025	0.084	0.283	0.324	-0.040	0.626	-0.073	0.288	0.764
Peer Groups									
	Men/Women			Men only			Women only		
Statistic	L	A	G	L	A	G	L	A	G
Mean	0.334	0.355	0.311	0.360	0.346	0.294	0.307	0.355	0.338
Stdev	0.033	0.025	0.030	0.027	0.020	0.024	0.028	0.022	0.022
CV (%)	9.74	7.060	9.77	7.42	5.91	8.20	9.26	6.22	6.61
Skewness	0.842	-0.553	0.520	0.622	-0.925	0.297	-0.431	-0.051	0.937

Table 8. Friedman test for peer groups and individual households

	Peer groups vs individuals					Individuals					Peer groups				
Source	SS	df	MS	χ^2	$p > \chi^2$	SS	df	MS	χ^2	$p > \chi^2$	SS	df	MS	χ^2	$p > \chi^2$
Columns	45.83	5	9.17	13.1	0.023	0.67	2	0.33	0.67	0.7165	4.67	2	2.33	4.67	0.097
Error	6.67	10	0.67			5.33	4	1.33			1.33	4	0.33		
Total	52.50	17				6.0	8				6.0	8			

5. Conclusions

Peer groups outperforms conventional individuals as the former has lower credit risk as compared to the latter. The diversification of risk in the peer groups inherent in individuals motivates a need to develop a mobile micro credit facility to cater for peer groups of the low income earners in Kenya. The credit scores and credit quality levels among the low income earners is higher in women than in men. When both men and women are analyzed as one, the overall scores are higher than those of men only peer groups but lower than those of the women. Financial products where both men and women are involved or women only peer groups would be more stable and can offer the providers lower credit risk levels.

Lending to women offers lower credit risk as compared to men even for low income earners. Even though men have higher financial muscle than women due to the societal set up in our country, women offers a more stable lending environment due to their strong horizontal social networks. This can be explained by the fact that women borrow more from informal financial services. The women peer groups are more stable, with less variability in exhibiting credit scores as compared to men. The clustering pattern observable in women peer groups exhibits the high levels of social collateral they posses and the informal lending patterns and borrowing that characterizes them. Therefore, incorporating soft data already in use by M-Shwari to offer mobile based loans to peer groups and HMM is ideal for risk classification.

Further research can consider analyzing the credit risk profile of each peer group member and develop a mechanism to reward more consistent savers in the envisioned mobile micro credit system. On the same note, the system can consider how to estimate the moral hazard of the peer group members and cater for it, and allow for peer selection of the peer group membership.

References

[1] Jack, W., and Suri, T. (2011). Mobile money: The economics of M-Pesa. NBER Working Paper 16721.

[2] Mbiti, I., and Weil, D.N. (2011). Mobile banking: The impact of M-Pesa in Kenya. Brown University, Department of Economics, No. 2011-13.

[3] Cook, T., and Mckay, C. (2015). How M-Shwari works: The story so far. Access to Finance Forum No. 10 (Reports by CGAP and FSD).

[4] Financial Sector Deepening (FSD). (2014a). Shilingi kwa shilingi: The financial lives of the poor. Kenya Financial Diaries, Nairobi, Kenya.

[5] Financial Sector Deepening (FSD). (2014b). A buck short: What financial diaries tell us about building financial services that matter to low income women. Report by Bankable Frontier Associates, Nairobi, Kenya.

[6] Financial Sector Deepening (FSD). (2016). 2016 FinAccess household survey. Collaboration between Central Bank of Kenya, Financial Sector Deepening and Kenya National Bureau of Statistics, Nairobi, Kenya.

[7] Financial Sector Deepening (FSD). (2013). FinAccess national survey 2013: Profiling developments in financial access and usage in Kenya, Nairobi, Kenya.

[8] Simtowe, F., Zeller, M., and Phiri, A. (2006). Determinants of moral hazard in Microfinance: Empirical evidence from joint liability lending programs in Malawi. African Review of Money Finance and Banking, 5-38.

[9] Holmstrom, B. (1979). Moral hazard and observability. The Bell Journal of Economics, 10 (1): 79-91.

[10] Allen, F., Demirguc-Kunt, A., Klapper, L., and Peria, M. (2016). The foundations of financial inclusion: Understanding ownership and use of formal accounts. Journal of Financial Intermediation, 27: 130.

[11] Demirguc-Kunt, A., Klapper L., Singer, D., and Oudheusden, P. (2015). The global findex database 2014: Measuring financial inclusion around the World. Development Research Group Finance and Private Sector Development Team. Policy Research Working Paper 7255,The World Bank.

[12] Demirguc-Kunt, A., Klapper, L., Singer, D., Ansar, S,. and Hess, J. (2017). The global findex database: Measuring financial inclusion and the FinTech revolution 2017. International Bank for Reconstruction and Development. The World Bank.

[13] PriceWaterHouseCooper (PWC) (2016). DeNovo Q2 2016 FinTech recap and funding review and the un(der) banked in fintech's largest opportunity. DeNovo: A platform to understand how disruption impacts business strategy and what actions to take.

[14] Ivatury, G., and Pickens, M. (2006). Mobile phone banking and low income customers: Evidence from South Africa. Consultative group to assist the poor. World Bank, Washington.

[15] Ahlin, C., and Jiang, N. (2008). Can micro-credit bring development? Journal of Development Economics, 86 (1): 1-21.

[16] Chavan, P., and Ramakumar, R. (2002). Micro-credit and rural poverty: An analysis of empirical evidence. Economics and Political Weekly, 37 (10): 955-965.

[17] International Finance Corporation (IFC). (2017).How Fintech is reaching the poor in Africa and Asia: A start-up perspective. International Finance Corporation. World Bank Group. Note 34, March 2017.

[18] Jagtiani, J., and Lemieux, C. (2017). Fintech lending: Financial inclusion, risk pricing and alternative information. Research Department, Federal Reserve Bank of Philadelphia, Working Paper 17-17.

[19] Kranz, N.C. (2005). Indicators of financial access: Household level survey. The World Bank, Financial sector Vice-Presidency. Conference on measuring access held in Washington DC, October 2004; http: //we.worldbank.org.

[20] Rabiner, L.R. (1989). A tutorial on hidden Markov model and selected applications in speech recognition. Proceedings of the IEEE, 77(2).

[21] Bilmes, J.A. (2006). What hidden Markov models can do. IEICE Transactions, Information and Systems (Special issue on IEICE Transactions), E89-D (3).

[22] Hassan, R., and Nath, B. (2005). Stock market forecasting using hidden Markov model: A new approach. Proceedings of the 2005 5th International Conference on Intelligent Systems Design and Applications (ISDA).

[23] Netzer, O., and Srinivasan, J.M.L. (2008). A hidden Markov model of customer relationship dynamics. Marketing Science, 27 (2): 185-204.

[24] Crowder, M., Davis, M., and Giampieri, G. (2005). A hidden Markov model of default interaction. Quantitative Finance, 5 (2): 27-34.

[25] Srivastava, A., Kundu, A., Sural, S., and Majumdar, A.K. (2008). Credit card fraud detection using hidden Markov model. IEEE Transactions on Dependable and Secure Computing, 5 (1).

[26] Quirini, L., and Vannucci, L. (2014). Creditworthiness dynamics and hidden Markov models. Journal of the Operational Research Society, 65: 323-330.

[27] Ntwiga, D.B. (2016). Social network analysis for credit risk modeling. Unpublished PhD thesis, School of Mathematics, University of Nairobi, Kenya.

[28] Ntwiga, D.B., Ogutu, C., and Kirumbu, M.K. (2018) Inclusion of peer group and individual low-income earners in M-Shwari micro-credit lending: A hidden Markov model approach. International Journal of Electronic Finance, 9(2): 121-133.

[29] Daniel, B., and Grissen, D. (2015). Behavior revealed in mobile phone usage predict loan repayment. Social Science Research Network (SSRN), 2611775: 1-10.

[30] Marrez, H., and Schmit, M. (2009). Credit risk analysis in micro credit: How does gender matter? CEB Working paper No. 09/053, Universite Libre Bruxelles.

[31] Bhole, B., and Ogden, S. (2010). Group lending and individual lending with strategic default. Journal of Development Economics, 91: 348-363.

[32] Gomez, R., and Santor, E. (2003). Do peer group members outperform individual borrowers? A test of peer group lending using Canadian micro credit data. Bank of Canada working paper 2003-33.

[33] Financial Sector Deepening (FSD). (2014c). Kenya financial diaries datasets user guide. A guide to the data produced by the 2012-2013 Kenya financial diaries, Nairobi, Kenya.