**Statistical Modeling and Analysis of Access to Financial Services in Indonesia**

Multivariate Statistics Final Report

Submitted by: Saba Mundlay, Jonathan Vega, Szábo Sándor

April 8, 2016

**Table of Contents**

|  |  |
| --- | --- |
| Introduction | ………………………………………3 |
| Debt History and Credit Markets | ………………………………………4 |
| Labor Dynamics and Credit Markets | ………………………………………11 |
| Education and Credit Markets | ………………………………………17 |
| Conclusion | ………………………………………23 |
| Appendix A | ………………………………………23 |
| Appendix B | ………………………………………23 |

**Introduction**

A significant part of the discourse surrounding economic development in transitioning countries (such as Indonesia) is about access to credit: the ability to borrow money. In rural areas, formal financial institutions (public and private) aren’t as developed. Thus, communities rely on informal financial structures, such as community based savings schemes. Our project hopes to contribute to our body of knowledge on community structures and household dynamics and their respective influences on credit markets. This project can be seen as a step towards creating a relevant framework for assessing the effectiveness of economic development policies (identifying the various factors affecting rural use of credit markets). We use data from an Indonesian national survey conducted by The World Bank assessing access to financial services. Given that “the dominant reason for not having a bank savings account was mainly economic” (The World Bank, 2009), this project makes an effort to separate the intrinsically complex socio-economic factors that make up communities.

The survey was conducted using a multi-stage random sampling technique. This is a population-weighted selection method first at the provincial level, then the sub-province level and lastly at the village level. On the main island of Java, the final sample contained 34 urban villages and 30 rural villages. Off the main island, it contained 18 urban villages and 30 rural villages, due to the rural nature of being off the main island. Ultimately, 30 households were randomly selected in each village resulting in a final sample of 3,360 household respondents (The World Bank, 2009). While this is not a true random sampling method, it is the most effective for a dense financial survey. Thus, we believe that the survey is representative of the nation of Indonesia.

The survey finds that only half of Indonesia’s population has access to formal financial services. Physical access to formal financial services is mainly a rural problem. “Generally, households used bank, community welfare scheme or even arisan to save, but informal financial service providers were the most common to get a loan” (The World Bank, 2009).

**Debt History and Credit Markets**

Research Question:

How does the debt history of a household affect its participation in an informal community-based savings program (CBS)?

Brief Literature Review & Contribution Made:

A paper by Siwan Anderson, Jean-Marie Baland and Karl Over Moene on “Enforcement in Informal Savings Groups” finds that informal savings groups often rely on a social sanctioning. This brings into question a household’s trustworthy-ness by its neighbors. Much of this project is based on a paper titled “Social Networks and Credit Access in Indonesia” by Cagla Otken of Bilken University and Una Okonkwo Osili of Indiana University. They state that informal credit markets rely more heavily on third party guarantees, reputation and threat of loss of access to future credit than their formal counterparts. They find that “household debt has a negative and significant effect on obtaining credit (Otken, Osili, 2004: 11).” These results imply that the lender (the CBS) is able to observe the household’s debt position. This concept is further developed in a paper on the effects of cooperative credit structures by Abhijit Banerjee, Timothy Besley and Timothy Guinnane. They call this a “peer monitoring” view of debt because neighbors have better information about borrowers than formal banks and can use this information to influence their credit transactions with said borrower. Their paper concludes with three main reasons why credit cooperatives might function better in transitioning economies: (1) community social sanctions play a strong role in sustaining “non opportunistic” behavior amongst individuals (2) cooperatives are sustained by repeated interactions meaning shortsighted privately optimal behavior is curtailed and lastly that (3) cooperative credit structures function through monitoring. These sources serve as justification for exploring a relationship between debt history and participation in a CBS, given the transitioning nature of the Indonesian economy. Since the data used in the papers is different from our data set, they serve only as loose theoretical frameworks. Additionally, Otken and Osili also find that an interaction variable between participation in informal savings schemes and per capita expenditure of a household have a significant negative access on being granted credit. When asked whether there was access to an overdraft line of credit, 1120 respondents said “no” and only 87 said “yes.” Table 1 (below) shows the main reasons and their frequencies for not having a bank account. Note that the most popular reason was “do not have enough money (all spent for consumption).”

|  |  |
| --- | --- |
| Reason | Frequency Count |
| Bank fees too high | 22 |
| Bank location is not convenient | 40 |
| Bank staff are rude or unhelpful | 22 |
| Do not have a job | 162 |
| **Do not have enough money (all spent for consumption)** | **1710** |
| Do not see the advantage of having a bank account | 62 |
| Do not trust banks | 4 |
| Don’t know how the bank operates | 75 |
| Financial products are not suitable | 24 |
| Interest rate offered is too low | 2 |
| Prefer to invest money | 28 |
| Other | 2 |

Table 1

Data:

Two plausible relationships arise: (1) Current debt might prevent participation because households are not allowed to save, perhaps because of their expenditure habits. (2) A Household with a “bad” debt history might have a detrimental social reputation, which affects participation. Taking into consideration existing literature and the limitations of the data set, below are the chosen variables to operationalize the research question:

(1) past\_use\_credit: collapsed binary variable taken from the question “have you or any other member of your household every used (formal institutions, micro-finance institution, employer, pawn shop, daily bank, community welfare scheme, neighborhood community or credit from a shop) to borrow money?” If the respondent said yes to any of the above, I recoded the response to 1, else 0.

(2) spending\_habits: binary variable taken from the question “which of the following limit your household’s ability to save in any financial institution?” If the respondent said yes to “want to save, but have difficulty controlling my spending habits” I recoded it to 1, else 0.

(3) too\_much\_debt: binary variable taken from the question “which of the following limit your household’s ability to save in any financial institution?” If the respondent said yes to “we must pay off debts before we can save” I recoded it to 1, else 0.

(4) belief\_in\_savings: binary variable taken from the question “do you believe your household saves enough for the future?” If the respondent said yes I recoded it to 1, else 0.

(5) HH\_credit: binary variable taken from the question “was this house purchased using credit?” If the respondent said yes I recoded it to 1, else 0.

(6) community\_savings: this is the primary response variable. Participation in a CBS is as defined by The World Bank for the purposes of this survey. This includes savings in a community welfare program or other informal savings group (savings club, arisan, ROSCA, etc). It also includes participation in government programs intended to improve people’s welfare. Examples of this include: revolving fund, credit union village economic unit, etc. An informal savings group is one that is formed by the people and does not have any legal basis. This variable is a binary variable taken from the question “Have you or anyone in your household used a community welfare scheme or informal saving clubs (such as neighborhood rotating savings group) to save money in the past 12 months?” If the responded said yes I recoded it to 1, else 0.

Given the categorical nature of the variables, the data is analyzed with **log linear modeling,** ultimately finding three statistically significant models with varying degrees of substantive significance. Even though there is a clear response variables, it is important to note that log linear models do not show dependence, only association between the variables.

**Group 1: community\_savings, past\_use\_credit, belief\_in\_savings**

Log Linear Model: AB, BC, AC (Table 3)

At 𝝰 = 0.05, this model of no second-order interaction (NSOI) is the only one that fits this specific group of variables (refer to Appendix A for R script). The Likelihood Ratio and Pearson Value are relatively close, further supporting the fit of this model. Lastly, a check of the expected frequencies with the actual further show the fit of the model. The model indicates that interaction exists between every combination of two variables, but none between all three.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **X2** | **df** | **P ( > X2 )** |
| **Likelihood Ratio** | 1.08146 | 1 | 0.2983711 |
| **Pearson** | 1.07835 | 1 | 0.2990670 |

For a better understanding of the model, one can look at the odds ratios for the interactions. If a household has accessed credit in the past, it is 1.59 times more likely to have participated (or currently be participating in) a CBS. If a household believes it saves enough for the future, it is 1.34 times more likely to have participated (or currently be participating in) a CBS. These odds ratios indicate that the effect of having accessed credit in the past is slightly stronger than a household’s belief that it saves enough in the future. The substantive significance of these results is hard to determine because of two things. First, the odds ratios are barely above 1 meaning that the association between variables is weak. Second, the element of *time* is unclear*.* community\_savings and past\_use\_credit are both time independent because the questions ask “have you ever” or “are you currently.” It is possible that an individual only answered “yes” to community savings because that is how they had accessed credit in the past. belief\_in\_savings is time-dependent of when the question was asked, in this case it is tied to 2009. Thus, one cannot comment on what the household believed at the time of making the decision to participate or not in a CBS. This concern over time may violate one of the assumptions of Poisson distributed data, that occurrences are independent.

**Group 2: community\_savings, past\_use\_credit, too\_much\_debt**

Log Linear Model: AB, AC, BC (Table 4)

The statistical interpretation of this model is the same as for the one above. The odds ratios show that a household that has accessed credit in the past is 1.65 times more likely to have participated (or currently be participating) in a CBS. A household that has too much debt to start saving is 1.19 times more likely. This makes some substantive sense keeping in mind the idea that household with debt is less likely to be accepted by the community into an informal savings scheme, since it may not have the best reputation or best debt history.

Both Model 1 and Model 2 are of the same complexity, though Model 2 fits better. To see which model is more suitable, one can conduct an Akaike’s Information Criteria (AIC) test. This is necessary since the models have different groups of variables and are not nested. An AIC test takes into consideration the log likelihood ratio (to measure goodness of fit) and the number of parameters in the model (penalizing for more parameters). In this case, the latter category is irrelevant since the models have the same number of parameters. Below (Table 5) are the results of the AIC test. Since Model 2 has a slightly lower AIC, it statistically fits the data slightly better. It also seems to make more sense given the literature review. The model doesn't challenge any current discourse nor does it present particularly interesting information.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **X2** | **df** | **P ( > X2 )** |
| **Likelihood Ratio** | 0.4941800 | 1 | 0.4820686 |
| **Pearson** | 0.4419606 | 1 | 0.5061777 |

|  |  |
| --- | --- |
| **Model** | **AIC** |
| 1 | 15.08146 |
| 2 | 14.49418 |

**Group 3: community\_savings, spending\_habits, HH\_credit**

Log Linear Model (3): AB, AC (Table 6)

At 𝝰 = 0.05, this model of conditional independence was the only statistically significant fit for this specific group of variables. It is quite an interesting model to interpret. The odds ratios indicate that a household that was purchased using credit is 1.50 times more likely to have participated in (or be participating in) a CBS. A household that has difficulty controlling its spending habits is 1.22 times more likely. It makes sense that a household with difficulty controlling spending habits is less likely to participate in a CBS, given the “peer monitoring” idea expressed in the literature review. A model of conditional independence means that there is a relationship between HH\_credit and spending\_habits that can be “explained” away by community\_savings. This model indicates that the association between spending\_habits and community\_savings does not depend on HH\_credit, and the association between community\_savings and HH\_credit does not dependent on spending\_habits.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **X2** | **df** | **P ( > X2 )** |
| **Likelihood Ratio** | 0.8295604 | 2 | 0.6604854 |
| **Pearson** | 0.8378458 | 2 | 0.6577549 |

Limitations and Further Research:

None of the results above are particularly surprising, as they all fit within an existing body of knowledge. Moreover, the models are relatively complex and difficult to substantively comment on without relying on the literature review. At the most, I would use them as supporting evidence for the arguments made by the papers consulted. What is much more interesting is to look at the limitations of the data and the operationalization of the research question to shape further research questions.

One important limitation of the data was the survey structure. Probably for efficiency’s sake, not all household respondents answered every question. Thus, I had to pick and choose variables from questions that had been asked to the same respondents. Combined with making sure that the cell counts for each cross table were above five, it was difficult to come up with model for more than three variables at a time. More limitations arise when considered the structure of the analysis. First, there is a complete lack of contextual / control variables to account for any differences between households that may exist. Furthermore, no element of community infrastructure or leadership is directly taken into account.

A few interesting further research questions include looking at the different likelihood of participation between female heads of households vs. male heads of households. It would also be interesting to map out past spending and saving history according to other measures (such as quantity or frequency) and come up with a categorical scale to organize households into. Lastly, I would be most interested in creating a network graph to map out the social standing of various households to see how various structures of informal community-savings associations could spread throughout the network.

References:

Anderson, Siwan., Baland, Jean-Marie., Moene, Karl Ove., “Enforcement in Informal Savings Groups.” Journal of Development Economics, 90 (2009): 14 - 23.

Banerjee, Abhijit V., Besley, Timothy., Timothy W. Guinnane., “Thy Neighbor’s Keeper: The Design of a Credit Cooperative with Theory and a Test.” The Quarterly Journal of Economics, 109, no. 2 (1994): 491 - 515.

Otken, Cagla., Osili, Una Okonkwo., “Social Networks and Credit Access in Indonesia.” World Development 32, No. 7 (2004): 1225 - 1246.