University of Pennsylvania
ScholarlyCommons

Wharton Research Scholars | Wharton School

2016

ImpactScore: A Novel Credit Score for Social Impact

Simon Sangmin Oh
Wharton, UPenn

Jade Pooreum Lee
Wharton, UPenn

April I. Meehl
Wharton, UPenn

Follow this and additional works at: http://repository.upenn.edu/wharton_research_scholars
Part of the Business Commons

Oh, Simon Sangmin; Lee, Jade Pooreum; and Meehl, April I., "ImpactScore: A Novel Credit Score for Social Impact" (2016). *Wharton Research Scholars*. 135.
http://repository.upenn.edu/wharton_research_scholars/135

This paper is posted at ScholarlyCommons. http://repository.upenn.edu/wharton_research_scholars/135
For more information, please contact repository@pobox.upenn.edu.

ImpactScore: A Novel Credit Score for Social Impact

Abstract

Socially motivated lenders pursue lending that considers both financial return and social good, yet they lack a systematic tool to incorporate such considerations into their decisions. This paper proposes the application of credit scoring mechanisms not only to the likelihood of default but also to the likelihood of happiness. Using the existing data on microcredit loan applicants in Bosnia and Herzegovina, we construct a full credit scoring model that involves the construction of outcome variables to accurately capture borrower's change in subjective well-being, the classification of input variables depending on the ease of information acquisition, and the selection of the model based on different criteria. We also find that the variables on the household's level of consumption have significant explanatory power in predicting future subjective well-being of loan applicants.

Keywords

socially motivated lenders, credit scoring, subjective well-being, social finance

Disciplines

Business

This thesis or dissertation is available at ScholarlyCommons: http://repository.upenn.edu/wharton_research_scholars/135

ImpactScore: A Novel Credit Score for Social Impact

Pooreum Lee
The Wharton School
University of Pennsylvania
pooreum@wharton.upenn.edu

April I. Meehl
The Wharton School
University of Pennsylvania
meehlapr@wharton.upenn.edu

Sangmin Oh[1]
The Wharton School, School of Engineering and Applied Science
University of Pennsylvania
sangmino@wharton.upenn.edu

Jeremy Tobacman
Assistant Professor
Department of Business Economics and Public Policy
The Wharton School
University of Pennsylvania
tobacman@wharton.upenn.edu

[1] We would like to thank April and Simon's advisor, Lindy Black-Margida, for introducing us to the Wharton Research Scholars program as well as Dr. Utsav Schurmans and Professor Catherine Schrand for their assistance and support through the application process and year-long seminar. We would also like to thank Professor Jeremy Tobacman for his guidance and mentorship during the research process as well as for pushing us to become true researchers. Finally, we would like to extend our thanks to Professor Todd Gormley and Professor Paul Shaman for their advice on econometrics and overall research design.

Abstract

Socially motivated lenders pursue lending that considers both financial return and social good, yet they lack a systematic tool to incorporate such considerations into their decisions. This paper proposes the application of credit scoring mechanisms not only to the likelihood of default but also to the likelihood of happiness. Using the existing data on microcredit loan applicants in Bosnia and Herzegovina, we construct a full credit scoring model that involves the construction of outcome variables to accurately capture borrower's change in subjective well-being, the classification of input variables depending on the ease of information acquisition, and the selection of the model based on different criteria. We also find that the variables on the household's level of consumption have significant explanatory power in predicting future subjective well-being of loan applicants.

Keywords: socially motivated lenders, credit scoring, subjective well-being, social finance

1. INTRODUCTION

In a personal loan market, lending decisions are made through the collection and analysis of large amounts of data on variables that correlate with a borrower's probability of defaulting upon the loan. Traditional lenders utilize this information to develop a credit score that numerically predicts this likelihood of default and therefore, the expected financial return to the lender. Some lenders, however, are motivated by goals other than financial return and are instead concerned with the social impact of the loan.

These lenders, such as ethical banks, Community Development Finance Institutions (CDFI Fund), and Microfinance Institutions (MFIs), are interested in inherently different questions: how many jobs will be created from the loan? How will the loan contribute to the community? What is the impact on the environment? And most importantly, how happy will the borrowers be with the loan?

As these lenders are concerned with more than just financial return, traditional credit scores are not an appropriate tool for the lending decision process. Supplemental methods have been created, mostly for use of MFIs, that do combine financial and social concerns for lenders, but no current tool exists that does so through the use of statistical credit scoring techniques. A new scoring algorithm which applies traditional credit scoring mechanisms to both the likelihood of default as well as the likelihood of an increase in subjective well-being for the borrower would be better suited for these socially motivated lenders. This paper aims to prove the possibility of such an algorithm and constructs a basic model for estimating the borrower's increase in subjective well-being. The final product of this model, analogous to the traditional credit score, is the Impactscore.

We use publicly available data from “The Impacts of Microcredit: Evidence from Bosnia and Herzegovina” by Augsburg, De has, Harmgart, and Meghir (2015). We choose this dataset for various reasons. First, it contains both baseline and follow-up survey responses from loan borrowers, thus enabling a detailed panel study on their characteristics. Second, it focuses on individual loans instead of group loans, which matches our desired unit of study. Third, it contains information on the borrower’s delayed payment or default on the loan as well as the self-measured level of subjective well-being, which are critical outcome variables for our model. The study by Augsburg, De has, Harmgart and Meghir (2015) is thus ideal for our purpose and contains rich borrower characteristics including demographic details, spending patterns, and loan specifications.

The ImpactScore, the final output of our model, is based on two predicted probabilities: the probability of the borrower defaulting on a loan and the probability of borrower’s increase in subjective well-being. To arrive at the final output, we follow a three-step process. First, we construct outcome variables to accurately capture the borrower’s status due to the loan. Second, we select input variables to be used in the model and categorize them depending on the ease of information acquisition. Third, we choose the best model based on different criteria and generate the probabilities to be used in the lending decisions.

Specifically, the characteristics of borrowers and loans from the study are categorized into three sets, and the divisions are made based on the relative access that lenders would have to each piece of information. Each of these sets is then used to predict three different binary outcome variables: loan_default, SWB1, and SWB2. SWB1 is an indicator variable created to mark an increase in consumption of temptation goods of a given threshold as well as the creation of a new business. SWB2 indicates the decrease in stress by more than a given threshold, with stress being measured as a variable in the chosen dataset.

The probability of each of the outcome variables is estimated using ordinary least squares regression, logistic regression, probit regression, and penalized logistic regression, and the models are evaluated using criteria such as Kernel Density Estimation, ROC curves, Akaike Information Criterion, and true positive and false positive rates. The models are created to be flexible enough so that any lender could input their own thresholds in order to receive the most appropriate lending decisions for their specific goals.

Section 2 describes the related literature on credit scoring and subjective well-being. Section 3 explains the data used in our study, and Section 4 summarizes the overall methodology for our study. Section 5 discusses the results of our analysis, and Section 6 concludes.

2. RELATED LITERATURE

2.1 Credit Scoring

2.1.1 Design of Credit Score

Credit score design is in the interest of many lending organizations. While the actual formula for generating the credit score is unknown to the public, Thomas, Edelman, and Crook (1999) describe in detail the process involved in designing such score. There are three main categories of scorecards: statistical scorecard, judgmental scorecard, and hybrid scorecard. Statistical scorecards derive empirically relevant factors from data on past loans, whereas a judgmental scorecard is structured from expert judgment and institutional experience. The hybrid scorecard is a combination of the prior two.

The critical step in credit score design is the defining “bad loans.” A bad loan can be any loss-making client that, with perfect hindsight, the lender would have chosen to avoid. A precise, quantitative definition of “bad” is crucial for developing successful statistical models, and naturally a compilation of a list of client characteristics is necessary. Widely used characteristics include: demographics (gender, marital status, and education level), household information (years in residence, number of children), household assets (vehicles owned, ownership of residence) and financial flows (business revenue, monthly income, rent payment).

Different types of scoring are also recognized based on the outcome that is subject to prediction (Schreiner, 2001). Application scoring, for example, predicts the probability that a loan will turn “bad,” whereas behavioral scoring focuses on the probability that the *next* installment will be late. Also, collections scoring predicts the probability that a loan late for a given number of days will be late for another given number of days, and desertion scoring looks at the probability of a borrower applying for a new loan once the current loan is paid off.

2.1.2 Statistical Methods in Credit Scoring

Linear Discriminant Analysis (LDA), a popular classification technique originally developed by R. A. Fisher, has been widely used in credit scoring design. Its purpose is to find the discriminant function by maximizing the difference between two groups while the differences among the members of the same group are minimized. Among many applications of the technique, the first use of LDA is that of Durand (1941) who showed that the method produced reasonable estimates of credit repayment.

Logistic regression is also widely used. It involves calculating the log odds of a loan being “good” based off of a linear regression of multiple chosen variables. For a given loan being considered, the log odds can easily be rewritten as a percentage of a loan being “good,” and this likelihood can be compared to a pre-determined threshold for loan decision. This threshold is usually set by calculating the weighted misclassification error – the number of “good” loans classified as “bad” multiplied by the opportunity cost of not granting this loan added to the number of “bad” loans classified as “good” multiplied by the cost of default. As Schreiner (1999) points out, the perk of this approach is that although the regression model is created by the researcher, a lender can then choose the threshold based off of their own preference for risk.

The K-Nearest Neighbor (KNN) approach involves classifying an applicant as “good” or “bad” based on the proportion of “good” loans amongst the k nearest loans to the loan being studied. To use this approach, one must choose the distance metric. Often, it is typically chosen as a simple adaption to the typical Euclidean distance metric; Henely and Hand (1996) upgraded the approach by including the direction vector found in linear discrimination. Yet choosing the distance metric is of substantial complexity and the overall approach can be just as complicated as the regression-based approach to credit scoring. After determining the distance metric, one must choose the appropriate value of k and also the threshold for the minimum proportion of “goods” in the k nearest neighbors to classify the given loan as “good.” More specifically, it must be greater than the default cost of classifying a “bad” loan as “good” divided by the total costs from misclassification.

Recent papers employ more advanced techniques. For example, Kumar and Bhattacharya (2006) find that artificial neural network model comprehensively outperforms the LDA model in both training and test partitions of the data set. Some studies combine discriminant analysis with other models – Lee et al. (2002) argue that integrating backpropagation neural networks with traditional discriminant analysis improves the credit scoring accuracy. As is the case with any statistical modeling, the key objective is to find the balance between classification accuracy and computational efficiency.

2.1.3 Credit Scoring in Social Context

The first statistically derived credit scoring model for microfinance was created using logistic regression (Schreiner, 1999). The model was constructed using relatively inexpensive data, which serves as a significant improvement over traditionally used personal traits in loan decisions. Schreiner has also studied the social benefit that can come from microfinance loans – in one paper, he evaluates the worth, cost, depth, breadth, length, and scope of a microfinance institution in order to gain an accurate depiction of the welfare provided by the microfinance institution.

Since then, numerous credit scoring models for socially motivated lenders have been experimented – they utilize techniques such as Analytical Hierarchy Process (AHP) (Auoam, 2009), Fuzzy Analytical Hierarchy Process (FAHP) (Che et. al, 2010), Tobit Regression (Deininger and Liu, 2009; Sharma and Zeller, 1997; Zeller, 1998), Discriminant Analysis (Auoam et al., 2009; Diallo, 2006; Viganò, 1993) , Neural Networks (Blanco et al., 2013), Data Envelopment Analysis (Che et al., 2010), Logistic Regression (Dinh and Kleimeier, 2007; Kinda and Achonu, 2012; Shreiner, 1999; Van Gool et al., 2012), Multinomial Logistic Regression (Vogelgesang, 2003), Probit Regression (Reinke, 1998), or a combination of these techniques.

More complicated methods for credit scoring models include those similar to the Measuring Attractiveness by a Categorical Based Evaluation Technique, or MACBETH approach (De Corte et al., 2012). This approach, which is highly used in the public and private sectors, quantifies the degree of attractiveness of an attribute by comparing it to a designated "neutral" level of attraction and "good" level of attraction.

More recently, the working paper by Serrano-Cinca, Gutiérrez-Nieto, and Reyes (2013) uses the AHP to generate a credit score that also includes a measurement for social impact. According to our knowledge, this is the only paper that explicitly combines the probability of default and social impact to generate a single loan decision metric. In their paper, the authors quantify social impact based on six categories in the United Nations Millennium Development Goals: impact on employment, impact on education, equal opportunities, community outreach, impact on health, and impact on environment. The score is then calculated by weighting factors influencing the borrower's credit past, present, and future, with the social impact being factored into the future component.

2.2 Utility and Subjective Well-Being

2.2.1 Borrower Utility

In behavioral economics, the standard model of utility and concept of revealed preferences do not exactly apply. Rather, utility of an individual is divided into two types: decision utility and experienced utility. Decision utility refers to the utility incurred at the time of decision making while experienced utility refers to that measured while undergoing the experience or retrospectively after the experience has concluded (Kahneman 1997; Congdon, Kling, & Mullainathan 2011). In the microfinance realm, this division is especially applicable: for microcredit borrowers with little to no credit history, their expected utility at the time of taking up the loan may significantly differ from the actual utility they witness throughout the life of the loan.

Other scholars contribute further by identifying factors that influence and lead to inaccurate prediction of subjective well-being at the time of decision, such as predicted sense of purpose, perceived sense of control over one's life, family happiness, and social status (Benjamin et. al., 2012). Another explores the relationship between subjective well-being and economic growth and confirms that increase in income does not necessarily correlate with proportional increase in happiness (Stevenson, & Wolfers, 2008).

2.2.2 Measurement of Subjective Well-Being

There are two main approaches in assessing the impact of microcredit on happiness. The first approach looks at the self-reported levels of happiness from population surveys (Di Tella, MacCulloch, and Oswald, 2001; Becchetti and Conzo, 2010; Duflo, Banerjee, Glennerster, and Kinnan, 2013). For example, Di Tella, MacCulloch, and Oswald (2001) utilize the Euro-Barometer survey series containing information on individual happiness and life satisfaction level. Such information is very useful in forming the identification strategy of the research, but the associated measurement errors sometimes pose serious concerns.

The greatest benefit of self-reported subjective well-being measure is that the results are indeed subjective at an individual level. However, the use of respondents' evaluation about the quality of their life has inherent sources of error. For one, the signal of the inner state of the respondent may be impacted by the current state or temporary shocks exogenous to their ordinary lives. Another problem is that the ordinal scales across different cultures can be quite incomparable. A clear definition of happiness is also an area of continued debate, and defining which set of emotions to include could be a subjective task, depending on the given researcher choosing the emotions. Results can vary on the type of question: if, say, it is the amount of time that people experience positive affect that defines happiness, not necessarily the intensity of that affect, the results of self-reported happiness level can fluctuate on the duration that each question addresses (Lyubomirsky, King, Diener, 2005).

Another approach involves objective proxies of individual happiness levels (Mohindra, Haddad, and Narayana, 2008). Sometimes these proxies are preferred as they are more quantifiable and less prone to measurement error from surveys. The most frequently used proxies include changes in household income and assets, consumption of temptation goods, establishment of new business, and access to health services.

With enough historical data, identifying proxies with reliable predictability of subjective well-being, can reserve us statistical significance. One shortcoming of using proxies is that the results are not subjectively measured. Additionally, the representativeness of a synthetic indicator of borrower's life satisfaction in mirroring subjective well-being can vary greatly from population to population, which leaves the problem of incomparability unsolved.

2.2.3 Impact of Loans

We are primarily interested in loans that are likely to impact the borrower's livelihood and subjective well-being. The most prominent setting with such characteristics is that of a microloan, which is often used in regions with low-income families. As much as a microloan is issued with purpose of saving borrowers from social exclusion and financial disadvantages, happiness or self-esteem measure help quantify impact on the individual non-pecuniary benefit, and serve as a measuring stick in gauging overall performance of a microloan program in serving its borrowers.

Despite the many approaches, consensus is yet to be reached on the impact of microcredit on happiness. A group of studies finds no significant effect on prevalence of emotional stress or changes in life satisfaction (Ahmed, Chowdhury, and Bhuiya, 2001). A common concern for the finding is that the lack of significant effect may be due to the short period of microcredit interventions. Another concern is that the positive changes from increased income may be offset by emotional stress from additional liabilities. As Graham (2009) points out in her book, the mixed findings can be further attributed to the differences in population and choices of proxies for analysis.

Another group, on the other hand, documents significantly positive changes due to microcredit (Mohindra, Haddad, and Narayana, 2008; Fernald et al., 2008; Becchetti and Conzo, 2010). One channel of positive impact is improved healthcare access and the coverage of insurance costs; another is the increased consumption of goods that contribute to individual happiness. As indicated by Angelucci, Karlan, and Zinman (2015), the interpretation of the findings also hinges heavily on the proxies used to test different hypotheses.

The last group of researchers finds that microcredit may actually trigger depression and increased stress (Omorodion, 2007). The commonly provided rationale is that with increased access to credit, borrowers may be forced to take on additional burden related to work. Another argument, as indicated by Ahmed, Chowdhury, and Bhuiya (2001), is that many borrowers do not want to operate as entrepreneurs but are forced to do so due to loan specifications, thus experiencing an increase in stress.

3. DATA

To verify the efficacy of our model, we primarily rely on data publicly posted by academic publications. Many relevant research articles have been published by reputable economic journal publications, and a few of the data sets have been posted online. Primarily, we seek data sets that have both baseline and follow-up survey responses from the borrowers as well as questionnaires reflecting the borrower's status on the loan and changes in subjective well-being.

For our proof-of-concept, we utilize the data set from the paper "The Impacts of Microcredit: Evidence from Bosnia and Herzegovina" by Augsburg, De has, Harmgart, and Meghir (2015). Our initial candidates are from the January 2015 issue of the American Economic Journal, where six controlled experiments on impacts of microfinance programs are published. Among the six, only the study by Augsburg et al. (2015) fits our criteria; the others do not necessarily measure the impact of microfinance programs on individual participants or lack proxies of subjective well-being in their questionnaires.

Augsburg et al. (2015) analyzes the impacts of microcredit loans via randomized controlled trials on a group of marginalized loan applicants who have been previously rejected by a microfinance institution. The experiment takes place in Bosnia, and the data set contains both baseline and endline survey data that are rich in borrower characteristics, including demographic details, spending patterns, and loan characteristics. We find that this data set is the most complete out of all candidate data sets and thus ideal for our purpose of initial proof-of-concept.

[Insert Table 1 here]

More specifically, the authors identify a total of 1,241 marginal applications, of which 1,196 were approved and interviewed, and each applicant was allocated with a 50% probability to either the treatment (receiving a loan) or the control group (no loan). The baseline survey was conducted over the five-month period from February 2010 to July 2010, and 14 months after the participants were called back and invited to be re-interviewed. The attrition rate was approximately 17% with a 10 p.p. difference between the control and treatment group.

One important feature of this data set is their inclusion of survey questions on self-measured level of success. The survey contains 10 questions that measure various levels of anxiety, irritations, lack of control and confidence on a scale of 0 to 4 (0 = Never, 1 = Almost Never, 2 = Sometimes, 3 = Fairly Often, 4 = Very Often). The scores on each questionnaire were added to generate the variable *happiness_stress* which we ultimately use in our model.

[Insert Table 2 here]

Table 2 shows the descriptive statistics of the measured stress level per question. Each of the stress variables corresponds to a different survey question. The borrowers responded to these questions on a scale of 1 to 4, with 1 corresponding to never feeling the way described in the question and 4 corresponding to feeling said way very often. For all ten questions, the new microcredit did not seem to have a significant effect on the stress levels of the borrowers – the hypothesis that the difference between treatment and control is zero could not be rejected at the 5% significance level.

We also note that the means for each of these variables differs since some questions correspond to feelings often experienced while others represent feelings rarely felt. For example, *stress_difficulties* is the variable for a borrower's answer to "In the last month, how often have you felt difficulties were piling up so high that you could not overcome them?" As this is a very strong feeling, the mean for *stress_difficulties* is much lower than that of *stress_confidence*, the answer to "In the last month, how often have you felt confident about your ability to handle personal problems?"

[Insert Table 3 here]

Table 3 shows the descriptive statistics of the change in stress level for both the treatment and the control group. In addition to seeing no significant impact of the treatment, no significant change in stress was found between the means for stress for the baseline and endline surveys. The t-test for the change in the means of the aggregate of the stress levels between endline and baseline surveys showed a p-value of .3718, proving that there was no significant change.

[Insert Table 4 here]

Finally, for the purpose of this paper, we expand the data set by five to achieve a more stable model. Table 4 shows the descriptive statistics for *loan_default* for the expanded data set. For our predictive purposes, we are only interested in the treatment group – borrowers who were granted a micro loan in addition to their outstanding loans.

4. METHODOLOGY

The ImpactScore is created based on the two predicted probabilities: the probability of the borrower defaulting on a loan and the probability of the borrower's change in subjective well-being. We first describe the construction of different variables and then explain the selection process behind the dependent variables needed to estimate the probabilities.

4.1 Construction of Outcome Variables

The first outcome variable we are interested in is *default.* Specifically, we require information on whether the borrower has defaulted on the microloan. For our dataset, however, does not contain such information. Rather, it contains a variable *loan_default* which is equal to 1 if the borrower has ever defaulted on any of its loans, not only the micro loan. For our example, we use this variable to proxy for whether or not the borrower has defaulted on the current loan. This variable can also be thought of as representing general negative impact on the borrower's loan repayment ability.

The model also requires proxies for the borrower's current sense of well-being: happiness, life satisfaction, stress, and depreciation. Ideally, the dataset will contain information on all four variables, but our dataset only contains information on the borrower's stress level pre- and post-receiving of the loan. These variables are used to define an outcome variable that signifies change in subjective well-being.

For our purpose, we have created two custom subjective well-being variables, and they are summarized in Table 1. First, *SWB1* approximates the change in borrower's consumption of temptation as well as the fulfillment of their goal to own a business. More specifically:

$$SWB1_t = 1 \; if$$

$$1)\; business\_has_{t-1} = 0 \;\; \& \; business\_has_t \; = 1$$

$$2)\; \Delta consumption\_temptation_{t-1 \to t} \;\; \geq C_{threshold}$$

Note that $C_{threshold}$ can be determined by the individual lender. For our example, we use 10% for the threshold – in other words, *SWB1* is equal to one when the borrower who previously did not own a business started one during the period of the microloan *and* when the borrower's consumption of temptation goods increased by more than 10% during the period of the microloan.

Second, we define *swb2* as measure of the change in the borrower's stress level. Specifically,

$$SWB2_t = 1 \; if$$

$$\Delta happiness\_stress_{t-1 \to t} \; \leq C'_{threshold}$$

Note that $C'_{threshold}$ can also be determined by the individual lender. For our example, we use 10% for the threshold – in other words, *swb2* is equal to one when the borrower's self-assessed level of stress decreased by more than 10% during the period of the microloan.

4.2 Selection of Input Variables

The input variables required to construct the model need to be chosen with care. Typically, we consider the variables that are believed to be widely collected by lenders when deciding whether or not to grant a loan.

In this study, such variables are categorized into seven groups: Borrower, Consumption, Household, Business, Loan, Assets, and Subjective well-being. Variables in the Borrower category consist of those describing the borrower's status, such as level of education, age, and house ownership. Consumption contains the amount of money spent on goods such as clothing, food, and transportation. Household refers to the characteristics of the entire household and recent occurrences in it, such as crime, disasters, and deaths while Business applies to the current or new business managed by the borrower and its characteristics. Loan is used for the specific terms of past loans granted to the borrower, such as the interest rate, amount, and collateral. Assets is used for household ownership of vehicles, land, equipment, and other assets that are relevant to the household's wealth. Finally, Subjective well-being refers to the borrower's current sense of well-being, including measures for happiness, satisfaction, stress, and depression.

Although all of these variables are often collected in the determination of granting loans, it is likely that some lenders will not or will be unable to collect all of them. Therefore, we have split the variables into three sets: the restricted set, the medium set, and the expansive set. The Restricted set will include variables that majority of lending institutions definitely have accessible.

These include the variables found in the Borrower and Loan categories. The Medium set includes all variables in the Restricted set as well as the next set that lenders would be expected to collect, or the Household and Assets sets. Finally, the Expansive set contains all of the variables previously explained.

[Insert Table 5 here]

To account for the fact that data will not be available for many of these categories, we also create dummy variables for our analysis. These dummy variables are equal to zero if the lender has information for the corresponding input variable and one if the lender does not have the corresponding variable. In the event that a lender has collected most but not all variables of a given set of variables, the ImpactScore can still be run for that set of variables through the usage of the dummy variables.

While the introduction of additional groups of variables is expected to increase the accuracy, we avoid doing so for multiple reasons. First, we are restrained by the availability of data sets – only one of the six papers that we've examined contains a data set that fits our criterion. Also, as we want our design to be applicable to a large group of lenders, a more conservative design with the most widely used variables is recommended.

[Insert Table 6 here]

To avoid multicollinearity among the dependent variables, we examine the pairwise correlation matrix of the most important variables in our models. We find that the two most correlated variables are income from work and income from government with the correlation of $\rho = -0.2199$. Also, the level of consumption is positively correlated with both income from work and income from government.

4.3 Selection of Modeling Technique

To estimate the probability of default and change in subjective well-being, we utilize four different statistical techniques: OLS regression, logistic regression, probit regression, and penalized logistic regression. For each of the three outcome variables – *default*, *SWB1*, *swb2* – the four techniques are used using the three different sets of repressors – *restricted, medium, and expansive.* As a result, we obtain 12 different models and predictions for each of the given outcome variable.

4.3.1 OLS Regression

Using OLS regression to estimate a binary outcome is often referred to as a linear probability model. We essentially consider the following model:

$$Y_i = \boldsymbol{X\beta} + u_i$$

where N is the number of observations, K is the number of independent variables, $\boldsymbol{X}$ is the $K \times 1$ matrix of independent variables, and $\boldsymbol{\beta}$ is the $1 \times K$ matrix of coefficients. In this specification, β_k represents the change in probability of $Y = 1$ associated with a unit change in X_k. Thus, we have

$$p = \Pr(Y = 1 \mid \boldsymbol{X}) = \boldsymbol{X\beta}$$

An obvious problem with this approach is that the predicted values may not necessarily lie between 0 and 1. Probabilities must logically be between 0 and 1, but this model can predict probabilities outside this range.

4.3.2 Logistic Regression

Logistic regression is used to address predicted probabilities that lie outside [0, 1]. To do so, we make the following assumption:

$$p = P(Y = 1|\boldsymbol{X}) = \frac{\exp(\boldsymbol{X\beta})}{1 + \exp(\boldsymbol{X\beta})}$$

where Y is the binary response variable and $\boldsymbol{X} = [X_1, \dots, X_K]$ designate the explanatory variables. It thus follows that we can write:

$$\log\left(\frac{p}{1-p}\right) = \boldsymbol{X\beta}$$

4.3.3 Probit Regression

Probit regression is also used to address predicted probabilities that lie outside [0, 1]. Consider the following assumption:

$$p = \Phi(\boldsymbol{X\beta})$$

which implies that we are treating $\boldsymbol{X\beta}$ as a z-score. In other words, we can consider

$$Y^* = \boldsymbol{X\beta} + \epsilon$$

where $\epsilon \sim N(0, \sigma)$ with unknown σ. Then we can define

$$p = 1 \quad if\ Y^* > 0$$
$$p = 0 \quad if\ Y^* \leq 0$$

In this case, the probability can be derived as:

$$p = \Pr(Y = 1 \mid \boldsymbol{X}) = P(\boldsymbol{X\beta} + \epsilon > 0|\boldsymbol{X}) = P(\epsilon > -\boldsymbol{X\beta}|\boldsymbol{X})$$

$$= 1 - \Phi(-\boldsymbol{X\beta})$$

$$= \Phi(\boldsymbol{X\beta})$$

4.3.4 Penalized Logistic Regression

Penalized logistic regression is used to avoid overfitting of the model. Given the log likelihood function in a typical logistic model:

$$l(\boldsymbol{Y}, \boldsymbol{\beta}) = \sum_{i=1}^{N} Y_i \boldsymbol{X_i} \boldsymbol{\beta} - \log(1 + \exp(\boldsymbol{X_i} \boldsymbol{\beta}))$$

we add the penalization function $J(\boldsymbol{\beta})$ that discourages a high number of regressors. Thus the penalized negative log-likelihood is given as

$$-l(\boldsymbol{Y}, \boldsymbol{\beta}) + \frac{\lambda}{2} J(\boldsymbol{\beta})$$

The choice of λ is crucial and a procedure that estimates the optimal value of λ is needed. Also, a wide variety of penalty functions have been used, such as $\sum \gamma_k |\beta_k|$ and $\sum \gamma_k |\beta_k|^q$ $(0 < q < 1)$. To implement penalized logistic regression in Stata, we use a penalized logistic regression package *plogit* developed by Gareth Ambler at University College London. The penalization function used in this package is $\sum |\beta|$ which is equivalent to *Lasso*. We use $\lambda = 20$.

4.4 Validation

One of the main requirements for a good credit scoring model is high discriminatory power. There are many measures employed to assess the binary models – we propose the use of four most utilized criteria: kernel density estimation, Akaike Information Criterion (AIC), Receiver Operating Characteristic (ROC), and predictive power table.

4.4.1 Kernel Density Estimation

Kernel density estimation is a non-parametric way of estimating the probability distribution function (pdf) of a continuous random variable. For our purposes, it allows us to estimate the distribution of the predicted values from our model.

Conceptually, kernel estimators are similar to histogram but allow us to overcome the non-smoothness and dependence on end points that are inherent in histograms. Kernel estimators smooth the contribution of each observed data point over a local neighborhood of the data point, which is determined by the magnitude of the bandwidth. We first choose a kernel $K(u)$ which satisfies:

$$\int K(u)du = 1, K(u) \geq 0$$

We also denote the bandwidth as h. Then the estimated density at any point x is

$$\hat{f}(x) = \frac{1}{n}\sum K\left(\frac{x - x_i}{h}\right)$$

If the bandwidth h is too small, there is not much smoothing and leads to very spiky estimates; if h is too large, it leads to oversmoothing. We use the value of h that minimizes the Asymptotic Mean Integrated Squared Error (AMISE) assuming the data were Gaussian, which is the default metric in Stata.

4.4.2 Akaike Information Criterion (AIC)

Akiake Information Criteron (AIC) measures the relative quality of statistical models for a given set of data. It follows the following model:

$$AIC = 2k - 2\ln(L)$$

where L is the maximum value of the likelihood function and k is the number of estimated parameters in the model. The preferred model is the one with the minimum AIC value – it rewards goodness of fit but penalizes inclusion of more parameters. In the end, it is essentially penalizing overfitting of given data.

4.4.3 Receiver Operating Characteristic (ROC) & Predictive Power Table

A Receiver Operating Characteristic (ROC) curve plots the performance of a binary classification system as the discrimination threshold is varied. The curve is created by plotting the True Positive

(TP) rate against the False Positive (FP) rate. Generally, the closer the curve follows the left-hand border and then the top border of the graph, the more accurate is the classification. Conversely, the closer the curve comes to the 45-degree diagonal, the less accurate is the test.

A predictive power table illustrates a similar tradeoff between true positive and false positive but also provides a more granular overview of the classification accuracy.

5. RESULTS & DISCUSSION

In this section, we discuss the results and compare the models based on the four validation criteria. We first provide comparisons across the different scope of variables. This discussion is especially relevant because the variables that the lender can acquire varies significantly among regions, and thus identification of the most significant predictors greatly reduces the cost of information collection on the lender's part. We also provide comparisons of the power of different modeling techniques and their usefulness in classification. We focus on our subjective well-being outcome variables, *SWB1* and *SWB2.*

We first compare the classification results among using different scope of variables for model. Kernel density estimates provide us with a visual estimate of the classification: ideally, the two probability distributions would be significantly distinguishable from each other. First, we consider the case when *SWB1* is used as our outcome variable, which approximates the change in borrower's consumption of temptation as well as the fulfillment of their goal to own a business.

[Insert Figures 1 - 4 here]

Figures 1 ~ 4 contain the Kernel Density curves for SWB1 estimation across each variable scope and each modeling technique. For SWB1, we find that the restricted set of variables offers little predictive power in our model – the *pdf*s of those who are predicted to experience an increase in happiness ($Prob(swb1 = 1)$) and those who did not ($Prob(swb1 = 0)$) are not much distinguishable from each other. As we expand our regressors to the medium set, however, the distinction between the two distributions becomes much stronger. This pattern is consistent across all four modeling techniques. It is also interesting to note that expanding the regressors to the expansive set does not improve the visual classification as much.

[Insert Figures 5 - 8 here]

Figures 5 ~ 8 contain the Kernel Density curves for SWB2 estimation across each variable scope and each modeling technique. For SWB2, which is based on the borrower's self-reported level of stress, the pattern is slightly different: both the restricted set and the medium set of variable offer little predictive power in our model. In other words, the *pdf*s of those who are predicted to experience an increase in happiness ($Prob(swb2 = 1)$) and those who did not ($Prob(swb2 = 0)$) are not much distinguishable from each other. Only after we use the variables from the expansive set does the distinction between the two distributions become much stronger.

[Insert Figures 9 - 14 here]

We can also examine the ROC curves to visually assess the efficacy of our model. Figures 9 ~ 11 contain the ROC curves for SWB1 estimation and Figure 12 ~ 14 contain the ROC curves for SWB2 estimation. The visual pattern among the ROC curves are consistent with the kernel density estimates: for SWB1, expanding the variable set from restricted to medium significantly increases the discriminatory power; for SWB2, the expanding the variable set from medium to expansive increases the discriminatory power.

[Insert Table 7 here]

AIC and R-squared can also provide more quantitative measures of model quality. As a goodness-of-fit measure, AIC favors smaller residual errors but penalizes large number of predictors and potential overfitting. Table 7 provides the AIC values for each variable set. For both SWB1 and SWB2, expanding the variable set decreases the AIC value, indicating that the quality of the model increases with more inputs.

This finding is rather trivial – with more information about the borrower, we expect more accurate classification. What is of more importance is the change in AIC as we expand our variable set. For both SWB1 and SWB2, the decrease in AIC is larger when we expand our set from medium to expansive than from restricted to medium.

[Insert Table 8 here]

R-squared can also provide information about the explanatory power of our model. Table 8 provides the R-squared values, or pseudo R-squared values, for each variable set. The package used for penalized logistic regression does not report R-squared. The explanatory power increases slightly on average (2.27% to 7.67% for SWB1; 2.30% to 18.53% for SWB2) as we include more input variables in our model. It is interesting to note that the R-squared for SWB2 almost reaches 20%, whereas the R-squared for SWB1 is much smaller. One of the explanations for this asymmetry lies in the construction of our outcome variable SWB1. Because the binary variable is constructed based on two criteria (business fulfillment, consumption of goods), the model may not perform as well.

Finally, we examine the predictive power of each model. Tables provided in the online appendix illustrate the predictive powers for predicting SWB1. For the subjective well-being variables, we want to decrease the rate of people being classified as False Positives. These are people who are granted loans because they are expected to have increased subjective well-being from the loan, but who will actually have decreased subjective well-being, so it is very important to limit this rate. This is equal to 1 minus the True Negative Rate, therefore, we will look for thresholds that maximize the True Negative Rate. As the same time, we would like to decrease the number of False Negatives, or those who are not granted the loan but whose subjective well-being will actually increase from the loan.

For SWB1, thresholds increase with more variables, and the number of FN decreases (percentage change is large in each circumstance but the overall FN numbers are very smaller). FN numbers bigger across the board for Restricted, then smaller with each next scope. Therefore, with more information, the probability of swb1 = 1 actually decreases.

`Tables provided in the online appendix illustrate the predictive powers for predicting SWB2. More people are predicted to see decreases in happiness stress than those to see increases in consumption and fulfillment. Therefore, the thresholds we are considering need to be higher. Across the scopes, with more information, the probability of happiness stress decreasing is decreasing, with a greater decrease between restricted and medium than between medium and expansive.

Throughout our analysis, it was clear that regression and penalized logistic regression produced very similar results. True positive rates and true negative rates were very similar within each scope of variables, suggesting that the same thresholds could be chosen for these two techniques. Additionally, the results from logistic and probit regression were also almost exactly the same within each scope. The difference between the regression/penalized logistic regression results and the logit/probit results differs for each of the outcomes. Almost no difference is found amongst the probabilities for the four techniques when predicting swb1. For default, logit and probit have lower thresholds than regression and plogit while logit and probit have higher thresholds for swb2, both of which suggest that logit and probit predict lower probabilities for the outcomes than regression and penalized logistic regression do.

In addition, by studying the Kernel Density charts, we can see that within each scope, the distribution of predicted probabilities for each outcome does not vary much amongst the four techniques, just as was suggested by the predicted power tables. The only difference that is seen is that because OLS regression does not have a restriction in which predicted values must be greater than one, some of the values are less than one. However, amongst the predicted values that are greater than one, their distribution very closely matches those predicted through logit, probit, and penalized logistic regression for each outcome within each scope.

6. CONCLUSION

Socially motivated lenders, such as ethical banks and microfinance institutions, seek both financial return and social good. They are naturally interested in questions other than the likelihood of borrower repayment, and we have focused on the most challenging one: how happy will the borrowers be with the loan? Due to their goals, the lenders may need an alternate model to assess loan applications based not only on the projected profitability but also based on borrower benefits.

In essence, we have shown how credit scoring mechanisms can be applied not only to the likelihood of default but also to the likelihood of happiness. Using the data from the 2015 study of microcredit applicants in Bosnia and Herzegovina, we have constructed a model that involves the construction of outcome variables to accurately capture borrower's change in subjective well-being, the classification of input variables depending on the ease of information acquisition, and the selection of the model based on different criteria.

Our model can be flexibly adapted according to the client's needs. First, the outcome variable can be constructed depending on the lender's priorities and interest in different aspects of the borrower. Second, the input variables can be chosen depending on the borrower characteristics available to the lender. Finally, the classification tools can be replaced with more sophisticated techniques such as random forest or neural networks, if desired by the client.

Among the borrower characteristics used to predict future changes in subjective well-being, we have found the variables about the consumption level of households to be having significant explanatory power. As an extension of this research, it would be worthwhile examining which information on the consumption level is significantly related to future subjective well-being. This finding also has further implications on the type of information that lenders should seek to collect, and we hope further studies shed more light on the importance of such information.

7. REFERENCES

Ahmed, S. M., Chowdhury, M., & Bhuiya, A. (2001). Micro-Credit and Emotional Well-Being: Experience of Poor Rural Women from Matlab, Bangladesh. World Development, 29(11).

Angelucci, M., Karlan, D., & Zinman, J. (2015). Microcredit Impacts: Evidence from a Randomized Microcredit Program Placement Experiment by Compartamos Banco. American Economic Journal: Applied Economics, 7(1).

Aouam, T., Lamrani H., Aguenaou, S., Diabat, A. (2009). A Benchmark Based AHP Model for Credit Evaluation. International Journal of Applied Decision Sciences, 2(2).

Augsburg, B., De Haas, R., Harmgart, H., & Meghir C. (2015). The Impacts of Microcredit: Evidence from Bosnia and Herzegovina. American Economic Journal: Applied Economics, 7(1).

Bana e Costa, C. A., Decorte, J. M., & Vansnick, J. C. (2012). MACBETH. International Journal of Information Technology & Decision Making, 11(2).

Banerjee, Ab., Duflo, E., Glennerster, R., & Kinnan, C. (2013). The Miracle of Microfinance? Evidence from a Randomized Evaluation. American Economic Journal: Applied Science, 7(1).

Becchetti, L., & Conzo, P. (2010). Microfinance and Happiness. Facolta Di Economia Universit' Di Bologna, Sede Di Forli, Percorso di Studi in Economia Sociale.

Benjamin, D. J., Kimball, M. S., Heffetz, O., & Rees-Jones, A. (2012). What Do You Think Would Make You Happier? What Do You Think You Would Choose? American Economic Review, 102(5).

Blanco, A., Pino-Mejias, R., Lara, J., & Rayo, S. (2013). Credit Scoring Models for the Microfinance Industry Using Neural Networks: Evidence from Peru. Expert Systems with Applications, 40(1).

Che, Z. H., Wang, H. S., Chuang, C. (2010). A Fuzzy AHP and DEA Approach for Making Bank Loan Decisions for Small and Medium Enterprises in Taiwan. Expert Systems with Applications, 37.

Diallo, B. (2006). Un modele de 'credit scoring' pour une institution de microfinance Africaine: le cas de Nyesigiso au Mali. Laboratoire d'Economie d'Orleans (LEO), Universite d'Orleans.

Deininger, K., & Liu, Y. (2009). Determinants of Repayment Performance in Indian Micro-Credit Groups. Policy Research Working Paper 2885, Development Research Group, The World Bank.

Diener, E., Oishi, S., & Lucas, R. E. (2002). Subjective Well-Being: The Science of Happiness and Life Satisfaction. In C. R. Snyder & S. J. Lopez (Ed.), Handbook of Positive Psychology. Oxford and New York: Oxford University Press.

Dinh, T., & Kleimeier, S. (2007). Credit Scoring Model for Vietnam's Retail Banking Market. International Review of Financial Analysis, 5(16).

Di Tella, R., MacCulloch, R. J., & Oswald, A. J. (2001). The Macroeconomics of happiness. American Economic Review, 91(1).

Durand, D. (1941). Risk Elements in Consumer Installment Financing. Studies in Consumer Installment Financing: Study 9, National Bureau of Economic Research.

Fernald, L., Hamad, R., Karlan, D., Ozer, E., & Zinman, J. (2008). Small Individual Loans and Mental Health: A Rondomized Controlled Trial Among South African Adults. BMC Public Health, 8(409).

Henley, W. E., & Hand, D. J. (1996). A K-Nearest-Neighbor Classifier for Assessing Consumer Credit Risk. The Statistician, 45(1).

Kahneman, D., Wakker, P. P., & Sarin, R. (1997). Back to Bentham? Explorations of Experienced Utility. Quarterly Journal of Economics, 112.

Kahneman, D., & Krueger, A. B. (2006). Developments in the Measurement of Subjective Well-Being. Journal of Economic Perspectivs, 20(1).

Kinda, O., & Achonu, A. (2012). Building a Credit Scoring Model for the Savings and Credit Mutual of the Potou Zone. Consilience: The Journal of Sustainable Development, 7(1).

Kling, J. R., Congdon, S., & Mullainathan, S. (2011). Policy and Choice. Washington, DC: Brooking Institution Press.

Kumar, K., & Bhattacharya, S. (2006). Artificial Neural Network vs Linear Discriminant Analysis in Credit Ratings Forecast: A Comparative Study of Prediction Performances. Review of Accounting and Finance 5(3).

Lee, T., Ciu, C., Lu, C., & Chen, I. (2002). Credit Scoring Using the Hybrid Neural Discriminant Technique. Expert Systems with Applications, 23(3).

Lyubomirsky, S., King, L., & Diener, E. (2005). The Benefits of Frequent Positive Affect: Does Happiness Lead to Success? Psychological Bulletin, 131(6).

Mohindra, K. S., Haddad, S., & Narayana, D. (2008). Can Microcredit Help Improve the Health of Poor Women? Some Findings From a Cross-sectional Study in Kerala, India. International Journal of Equity Health, 7(2).

Omorodion, F. I. (2007). Rural Women’s Experiences of Microcredit Schemes in Nigeria: Case Study of Esan Women. Journal of Asian and African Studies, 42(6).

Reinke, J. (1998). How to Lend Like Mad and Make a Profit: a Micro-credit Paradigm versus the Start-up Fund in South Africa. Journal of Development Studies, 34(3).

Schreiner, M. (1999). A Scoring Model of the Risk of Costly Arrears at a Microfinance Lender in Bolivia.

Schreiner, M. (2004). Benefits and Pitfalls of Statistical Scoring for Microfinance. Savings and Development, 28(1).

Serrano-Cinca, C., Gutierrez-Neito, B., & Reyes, N. M. (2013). A Social Approach to Microfinance Credit Scoring. Solvay Brussels School of Economics and Management Centre, Universite Libre de Bruxelles.

Sharma, M., & Zeller, M. (1997). Repayment Performance in Group-based Credit Programs in Bangladesh: An Emprical Analysis. World Development 25(10).

Stevenson, B., & Wolfers, J. (2008). Economic Growth and Subjective Well-Being: Reassessing the Easterlin Paradox. NBER Working Paper No, 14282, National Bureau of Economic Research.

Thomas, Lyn C., Edelman David B., & Crook Jonathan A. (1999). *Credit Scoring and Its Applications.* Philadelphia, PA: Society for Industrial and Applied Mathematics.

Van Gool, J., Verbeke, W., Sercu, P., & Baesens, B. (2012). Credit Scoring for Microfinance: Is It Worth It? International Journal of Finance and Economics, 17(2).

Vogelgesang U. (2003). Microfinance in Times of Crisis: the Effects of Competitoin, Rising Indebtedness, and Economic Crisis on Repayment Behavior. World Development, 31(12).

Vigano, L. (1993). A Credit Scoring Model for Development Banks: An African Case Study. Savings and Development, 17(4).

Zeller, M. (1998). Determinants of Repayment Performance in Credit Groups: the Role of Program Design, Intragroup Risk Pooling, and Social Cohesion. Economic Development and Cultural Change 46(3).

Table 1 - Description of the Variables Used

Variable Name	Variable Description	Description
general_baseline	Timing of Survey	Dummy Variable = 1 if response is from follow-up survey
borrower_age	Age	Age of the borrower in years
borrower_marital	Marital Status	Indicator Variable = 1 if respondent is married; 2 if separated; 3 if single
borrower_education	Education Level	Dummy Variable = 1 if respondent completed high school education
borrower_school	School Enrollment	Dummy Variable = 1 if respondent is currently in school
borrower_dwelling	Dwelling	Dummy Variable = 1 if respondent owns dwelling
consumption_clothes	Amount spent on clothing	Average monthly amount spent on clothing in local currency in the past year
consumption_school	Amount spent on education	Average monthly amount spent on education in local currency in the past year
consumption_furniture	Amount spent on furniture	Average monthly amount spent on furniture in local currency in the past year
consumption_appliance	Amount spent on appliances	Average monthly amount spent on appliances in local currency in the past year
consumption_vehicle	Amount spent on vehicles	Average monthly amount spent on purchase of vehicle in local currency in the past year
consumption_repair	Amount spent on repairs	Average monthly amount spent on repairs in local currency in the past year
consumption_combustible	Amount spent on combustibles	Average monthly amount spent on combustibles in local currency in the past year
consumption_temptation	Amount spent on temptation goods	Average monthly amount spent on temptation goods in local currency in the past year
consumption_transportation	Amount spent on transportation	Average monthly amount spent on transportation in local currency in the past year
consumption_news	Amount spent on news	Average monthly amount spent on newspapers and magazines in local currency in the past year
consumption_recreation	Amount spent on recreation	Average monthly amount spent on recreation in local currency in the past year
consumption_food	Amount spent on food	Average monthly amount spent on food in local currency in the past year
consumption_medical	Amount spent on medical treatment	Average monthly amount spent on medical expenses in local currency in the past year
household_incomework	Income from work	Average monthly income from work in local currency in the past year
household_incomegovernment	Income from government	Average monthly income from government in local currency in the past year
household_kids	Kids in household	Number of kids aged under 17 in the borrower's household
household_death	Death in household	Dummy Variable = 1 if respondent's household experienced a death in the past year
household_illness	Illness in household	Dummy Variable = 1 if respondent's household experienced an illness in the past year
household_doctorvisit	Doctor visit in household	Dummy Variable = 1 if respondent's household member visited doctor in the past year
household_jobloss	Job loss	Dummy Variable =1 if respondent's household member lost a job in the past year
household_crime	Crime	Dummy Variable =1 if respondent's household reported any incident of crime in the past year
household_disasters	Natural disaster	Dummy Variable = 1 if respondent's household experienced a natural disaster in the past year

household_harvest	Bad harvest	Dummy Variable = 1 if respondent's household experienced a bad harvest in the past year
business_hours	Hours on business	Average hours per month spent on business and enterprise in the past year
business_wageempl	Hours on wage employment	Average hours per month spent on wage employment in the past year
buseinss_has	Ownership of business	Dummy Variable =1 if the respondent's household owns a business at the time of response
business_revenue	Business revenue	Average monthly revenue from business in the past year
business_expense	Business expense	Average monthly expense from business in the past year
assets_house	Assets - house	Value of the owned house in local currency
assets_land	Assets - land	Value of the owned land in local currency
assets_vehicle	Assets - vehicle	Value of the owned vehicle in local currency
assets_animal	Assets - animal	Value of the owned animals in local currency
loan_amount	Amount of outstanding loans	Amount of existing loans from microfinance institutions
loan_num	Number of outstanding loans	Number of existing loans from microfinance institutions
loan_interest	Interest rate on outstanding loans	Average interest rate on existing loans from microfinance institutions
loan_collateral	Collateral for outstanding loans	Dummy Variable = 1 if collateral was provided for existing loans
loan_purpose	Purpose of outstanding loans	Dummy Variable = 1 if outstanding loans were used for business expenses
happiness_stress	Stress level	Raw score on the survey question on level of stress
happiness_satisfaction	Satisfaction level	Raw score on the survey question on level of satisfaction
happiness_depression	Depression level	Raw score on the survey question on level of depression
happiness_locus	Locus level	Raw score on the survey question on level of control

Table 2 - Questionnaires for Stress Variable

Variable name	Questionnaire Item
stress_upset	In the last month, how often have you been upset because of something that happened unexpectedly?
stress_control	In the last month, how often have you felt that you were unable to control the important things in your life?
stress_nervous	In the last month, how often have you felt nervous and "stressed"?
stress_confidence	In the last month, how often have you felt confident about your ability to handle your personal problems?
stress_flow	In the last month, how often have you felt that things were going your way?
stress_cope	In the last month, how often have you found that you could not cope with all the things that you had to do?
stress_irritations	In the last month, how often have you been able to control irritations in your life?
stress_control2	In the last month, how often have you felt that you were on top of things?
stress_control3	In the last month, how often have you been angered because of things that were outside of your control?
stress_difficulties	In the last month, how often have you felt difficulties were piling up so high that you could not overcome them?

* The answers were recorded on a scale of 0 to 4: 0 = Never, 1 = Almost Never, 2 = Sometimes, 3 = Fairly Often, and 4 = Very Often. The scores on each questionnaire were added to generate the happiness_stress variable

Table 3 - Descriptive Statistics (Stress Level per Question)

	Control Group		Treatment - Control	
	Mean	SD	Coeff.	p-value
stress_upset	1.276	1.094	-0.003	0.960
stress_control	0.778	1.039	-0.062	0.301
stress_nervous	1.230	1.096	-0.095	0.134
stress_confidence	3.515	0.744	0.019	0.672
stress_flow	3.099	0.831	-0.010	0.836
stress_cope	1.004	1.044	-0.087	0.153
stress_irritations	2.961	1.073	-0.057	0.370
stress_control2	3.330	0.735	-0.004	0.925
stress_control3	1.190	1.085	0.041	0.514
stress_difficulties	0.789	0.976	-0.014	0.810

* Table 3 illustrates the descriptive statistics of the answers to the survey questionnaires in the data set. We find no significant difference in the mean responses to the questions between the treatment and the control group.

Table 4 - Descriptive Statistics (Stress Level)

	Control Group		Treatment - Control	
	Mean	SD	Coeff.	p-value
Stress Level				
Baseline	18.971	4.070	-0.054	0.839
Endline	19.025	5.073	0.193	0.537
Change between baseline ~ endline (%)				
(Endline-Baseline) / Baseline	4.933	38.104	1.539	0.372

* Table 4 illustrates the descriptive statistics of the responses to the questionnaires related to level of stress. During the period of the survey, the respondents experience an average of 4.93% increase in stress level. The difference of the increase between the treatment and the control group, however, are insignificant.

Table 5 - Classification of Variables

Restricted	Medium	Expansive
Gender	Income from work	Amount spent on clothing
Age	Income from government	Amount spent on education
Marital Status	Kids in household	Amount spent on furniture
Education Level	Death in household	Amount spent on appliances
School Enrollment	Illness in household	Amount spent on vehicles
Dwelling	Doctor visit in household	Amount spent on repairs
Amount of outstanding loans	Job loss	Amount spent on combustibles
Number of outstanding loans	Crime	Amount spent on temptation goods
Interest rate on outstanding loans	Natural disaster	Amount spent on transportation
Collateral for outstanding loans	Bad harvest	Amount spent on news
Purpose of outstanding loans	Assets - house	Amount spent on recreation
	Assets - land	Amount spent on food
	Assets - vehicle	Amount spent on medical treatment
	Assets - animal	Stress level*
	Hours on business	Satisfaction level*
	Hours on wage employment	Depression level*
	Ownership of business	Locus level*
	Business revenue	
	Business expense	

* Table 5 denotes the classification of the borrower characteristics into restricted / medium / expansive sets based on the ease of information acquisition.

Table 6 - Pairwise correlation matrix of selected variables

	Age	Amount of outstanding loans	Income from work	Income from gov.	Hrs. on business	Business revenue	Amount spent (temptation)	Amount spent (recreation)	Amount spent (food)	Stress level
Age	1.0000									
Amount of outstanding loans	-0.0912	1.0000								
Income from work	-0.1319	-0.0143	1.0000							
Income from gov.	0.0095	0.0334	-0.2199	1.0000						
Hrs on business	-0.1032	-0.0061	0.0965	0.0632	1.0000					
Business revenue	0.0216	-0.0020	-0.0103	-0.0016	0.0015	1.0000				
Amount spent (temptation)	-0.0072	-0.0028	0.0200	0.0158	-0.0341	-0.0013	1.0000			
Amount spent (recreation)	-0.0890	0.0221	0.0855	0.0254	0.0417	-0.0056	0.0506	1.0000		
Amount spent (food)	-0.1674	0.0435	0.1323	0.0162	0.0242	-0.0079	0.0064	0.1327	1.0000	
Stress Level	-0.0572	0.0263	0.0072	-0.0109	-0.0677	0.0588	0.0154	-0.0053	0.0148	1.0000

* Table 6 illustrates the pairwise correlation matrix of selected variables. We find that the two most correlated variables are income from work and income from government with the correlation of -0.2199. Also, the level of consumption is positively correlated with both income from work and income from government.

Table 7 - AIC Values for SWB1 and SWB2 Estimation

	OLS	Logit	Probit	Plogit
Outcome variable: SWB1				
Restricted	-1891.7	594.6	594.1	601.1
Medium	-1892.0	578.2	574.6	614.5
Expansive	-1928.1	510.4	512.4	608.2
Outcome variable: SWB2				
Restricted	2940.3	2811.7	2811.2	2816.6
Medium	2899.9	2772.0	2771.2	2796.1
Expansive	2860.1	2708.5	2707.0	2743.1

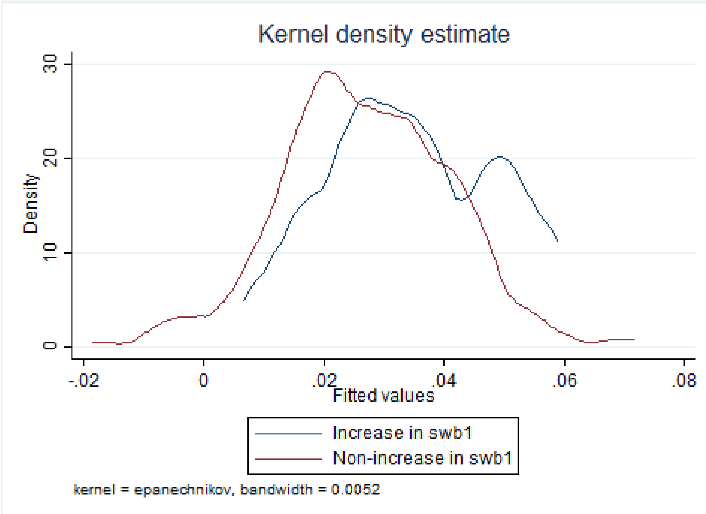
* Table 7 provides the AIC values for each variable set. For both SWB1 and SWB2, expanding the variable set decreases the AIC value, indicating that the quality of the model increases with more inputs.

Table 8 - R-squared Values for SWB1 and SWB2 Estimation

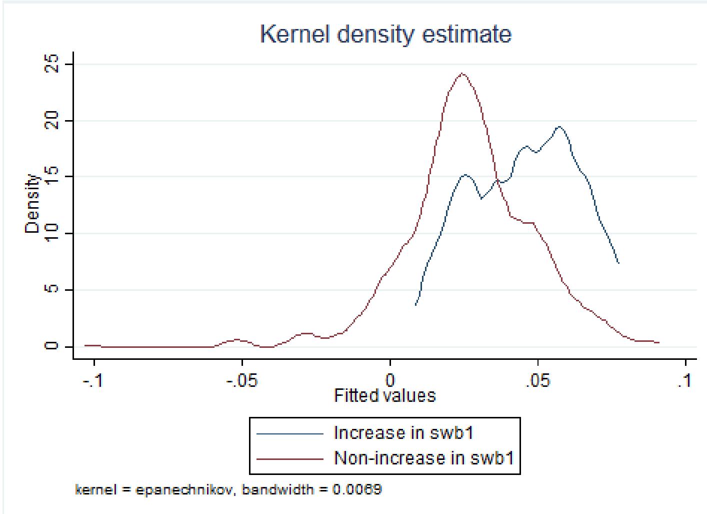
	OLS	Logit	Probit	Average
Outcome variable: SWB1				
Restricted	2.20%	2.30%	2.30%	2.27%
Medium	4.40%	4.60%	4.60%	4.53%
Expansive	7.70%	7.60%	7.70%	7.67%
Outcome variable: SWB2				
Restricted	0.80%	3.00%	3.10%	2.30%
Medium	1.80%	10.10%	10.70%	7.53%
Expansive	4.30%	25.80%	25.50%	18.53%

* Table 8 provides the R-squared values for each variable set. The package used for penalized logistic regression does not report R-squared. The explanatory power increases slightly on average as we include more input variables in our model. It is also interesting to note that the R-square for SWB2 almost reaches 20%, whereas the R-squared for SWB1 is much smaller.

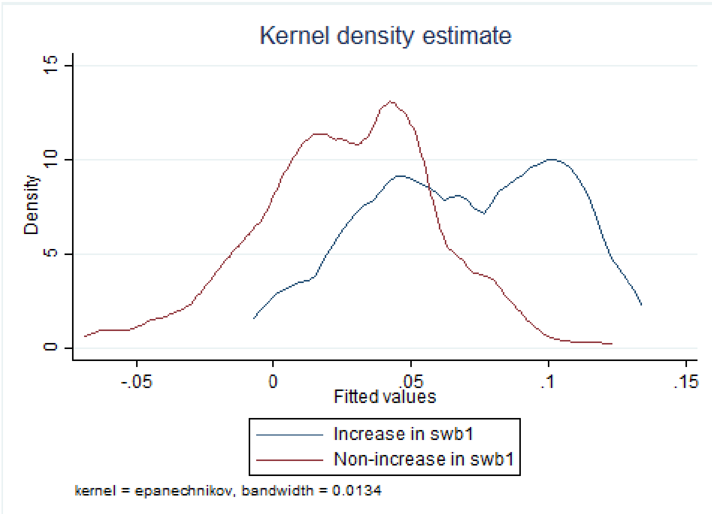
Figure 1 – Kernel Density Curve for SWB1 Estimation (OLS Regression)



(a) Restricted Set

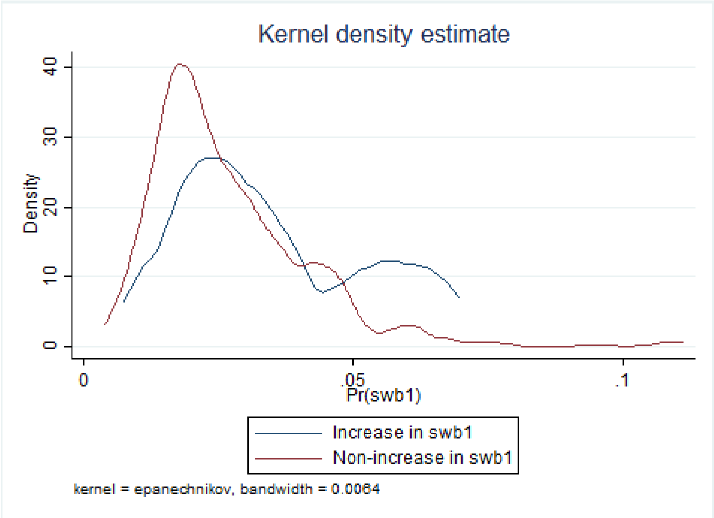


(b) Medium Set

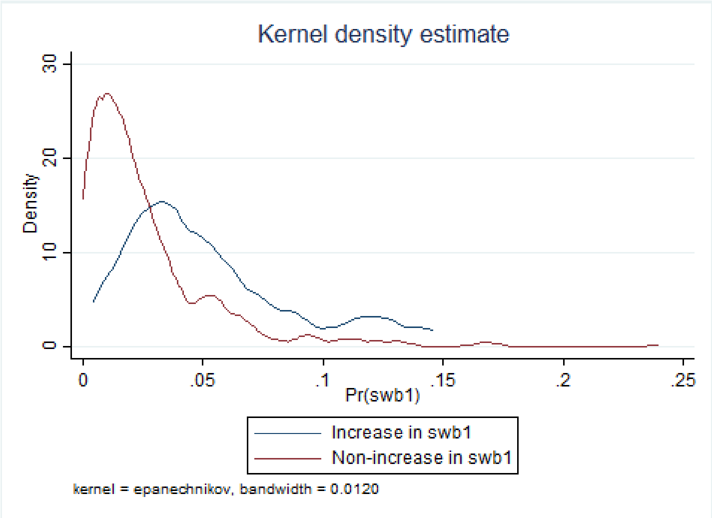


(c) Expansive Set

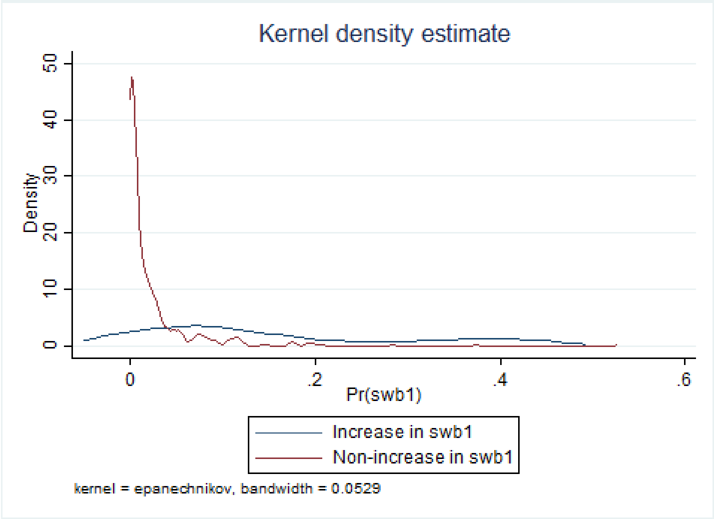
Figure 2 – Kernel Density Curve for SWB1 Estimation (Logistic Regression)



(a) Restricted Set

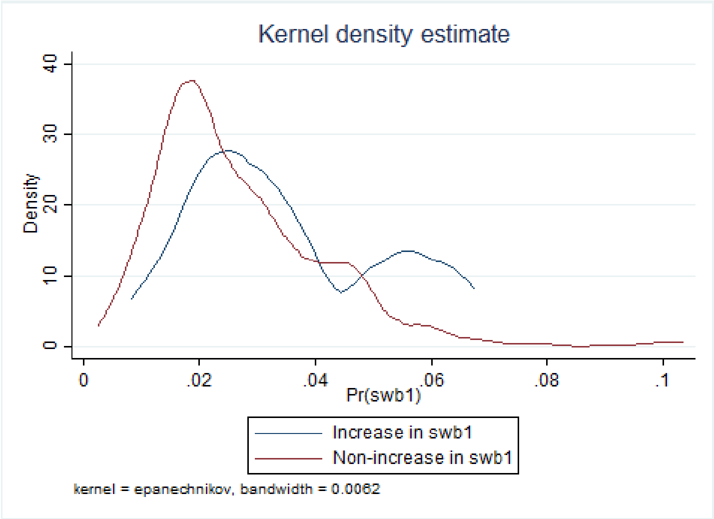


(b) Medium Set

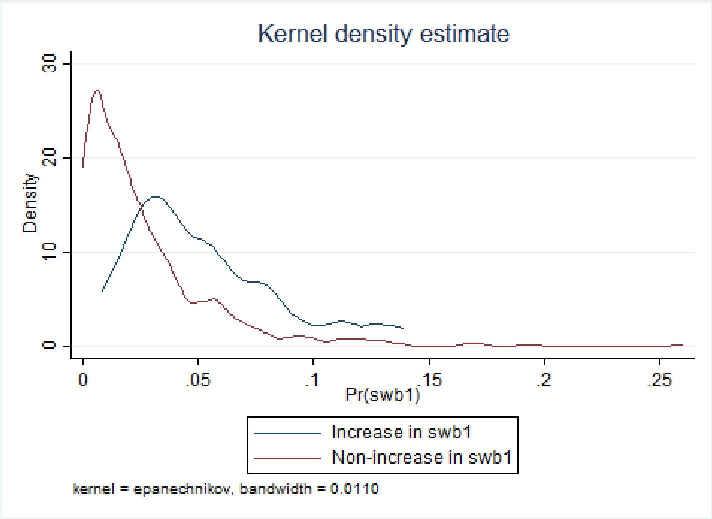


(c) Expansive Set

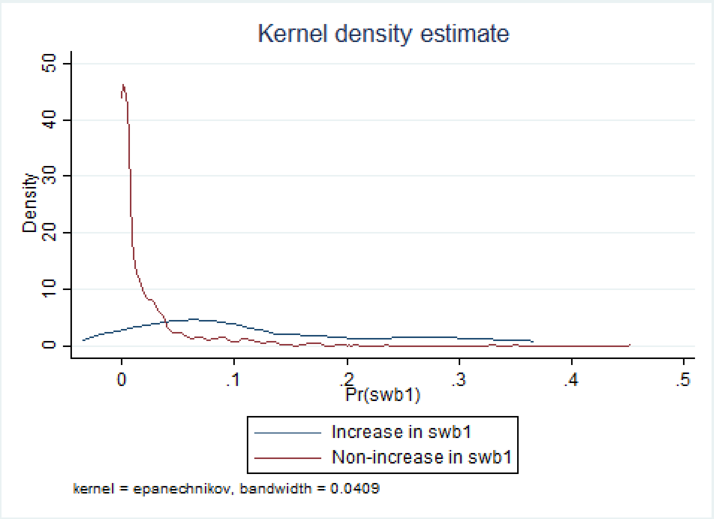
Figure 3 – Kernel Density Curve for SWB1 Estimation (Probit Regression)



(a) Restricted Set

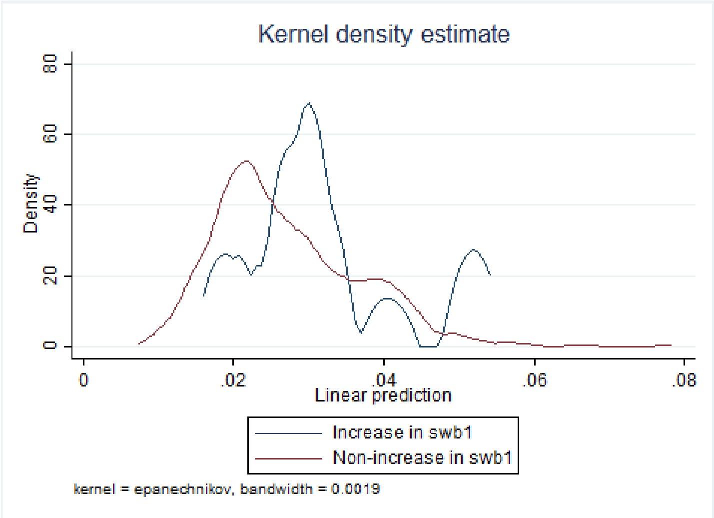


(b) Medium Set

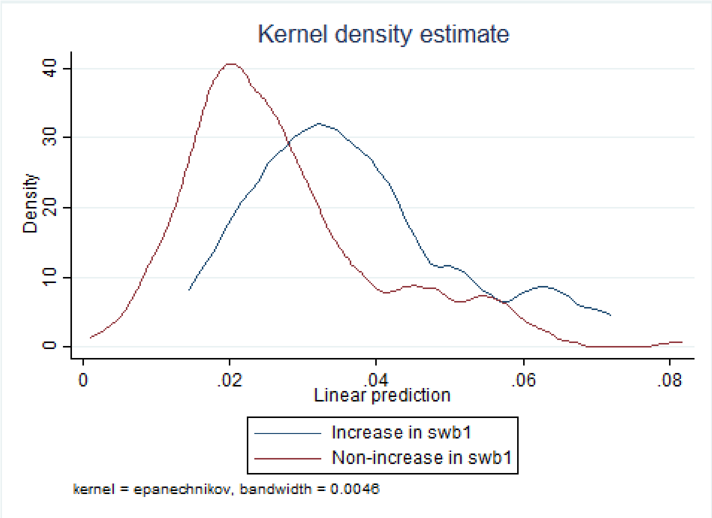


(c) Expansive Set

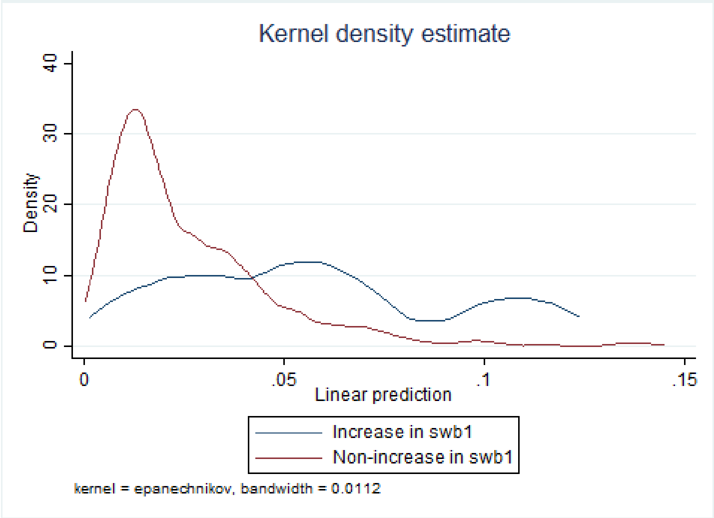
Figure 4 – Kernel Density Curve for SWB1 Estimation (Penalized Logistic Regression)



(a) Restricted Set

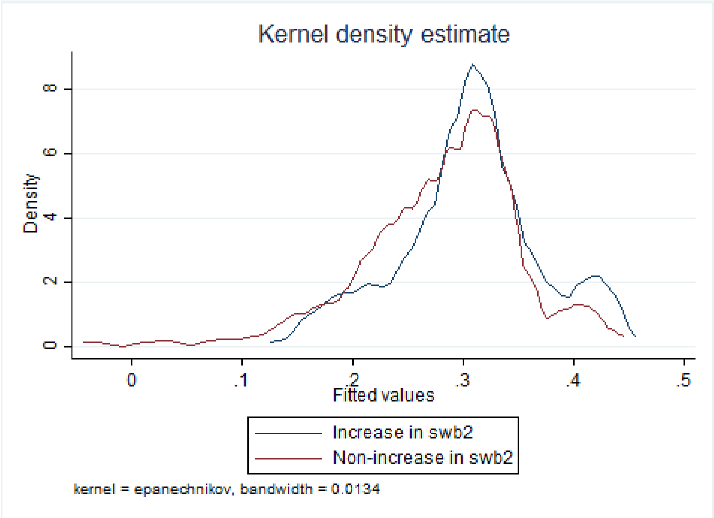


(b) Medium Set

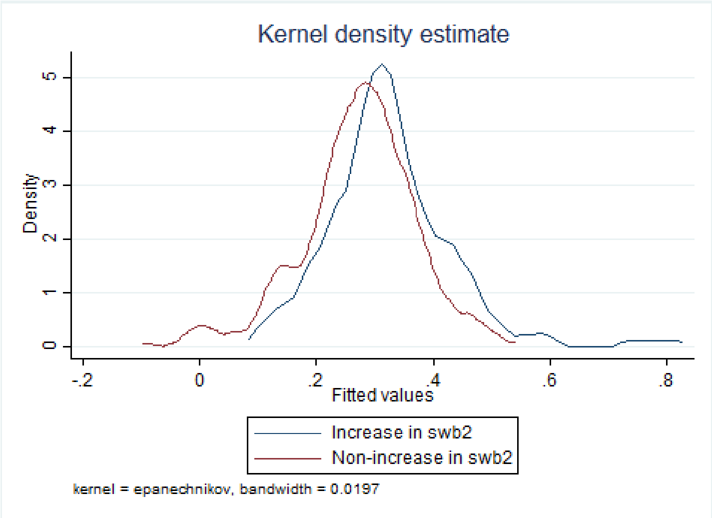


(c) Expansive Set

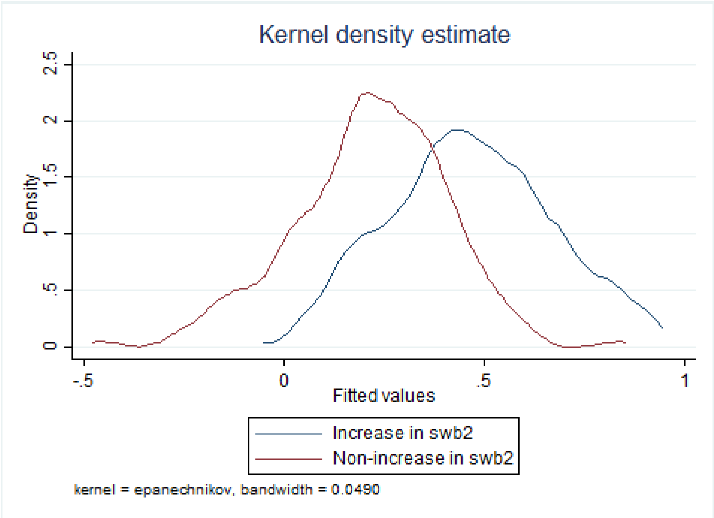
Figure 5 – Kernel Density Curve for SWB2 Estimation (OLS Regression)



(a) Restricted Set

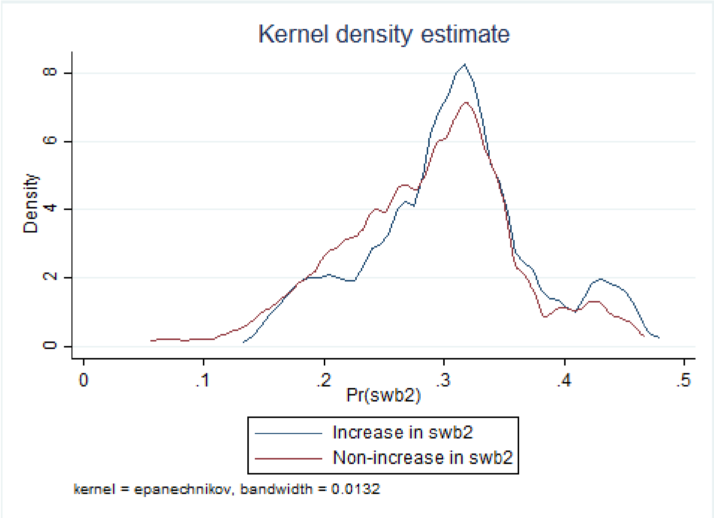


(b) Medium Set

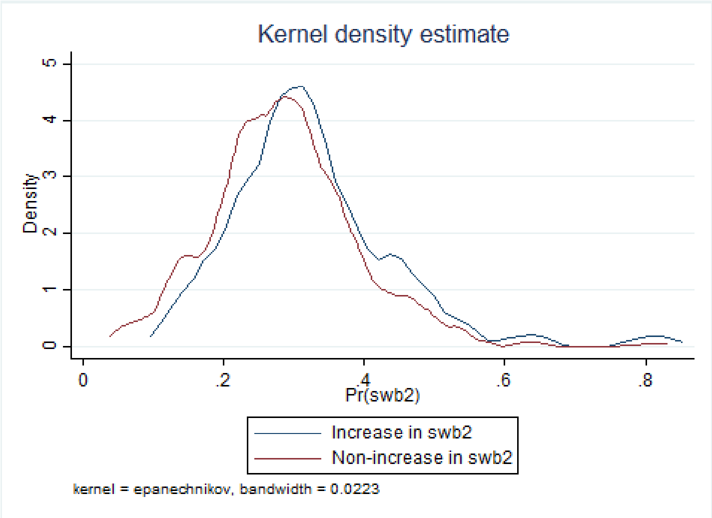


(c) Expansive Set

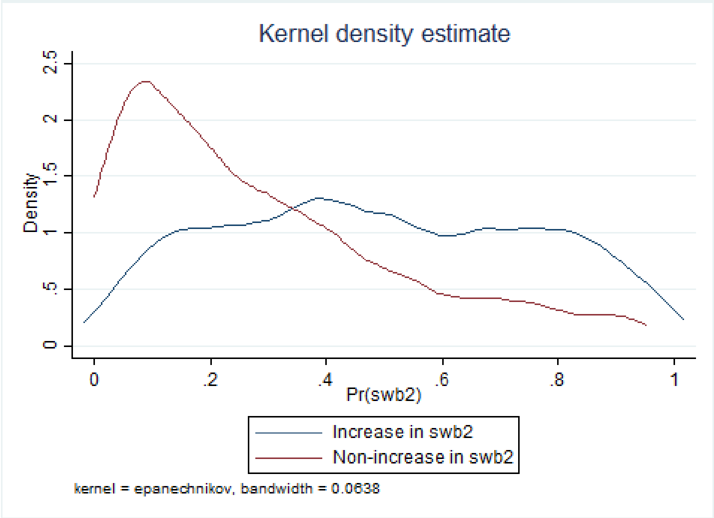
Figure 6 – Kernel Density Curve for SWB2 Estimation (Logistic Regression)



(a) Restricted Set

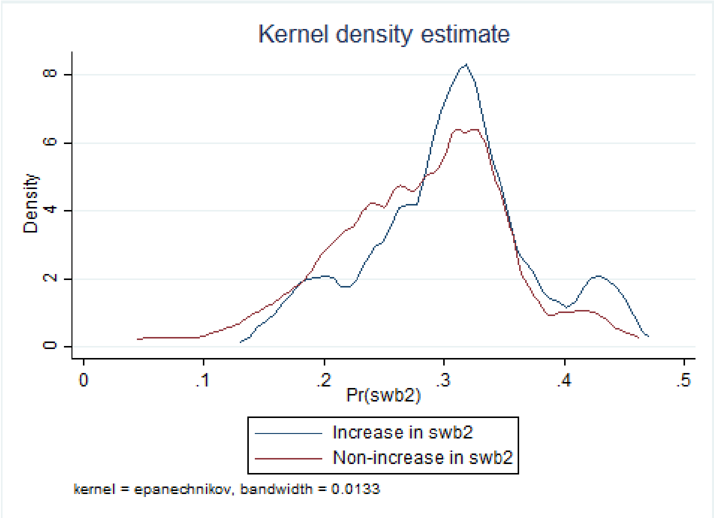


(b) Medium Set

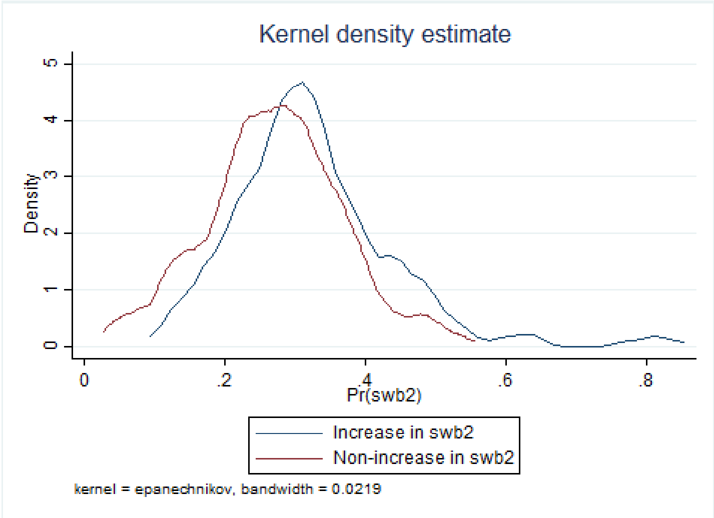


(c) Expansive Set

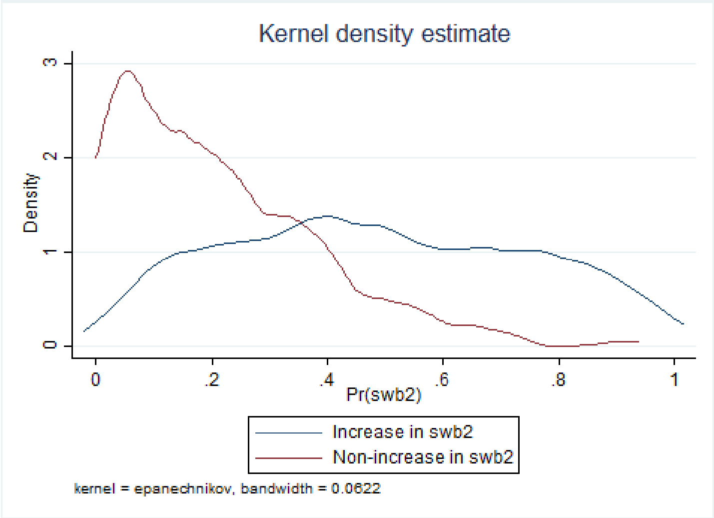
Figure 7 – Kernel Density Curve for SWB2 Estimation (Probit Regression)



(a) Restricted Set

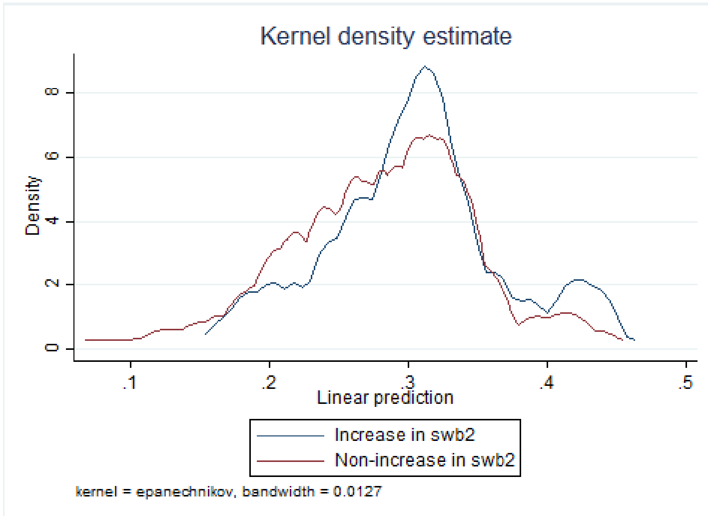


(b) Medium Set

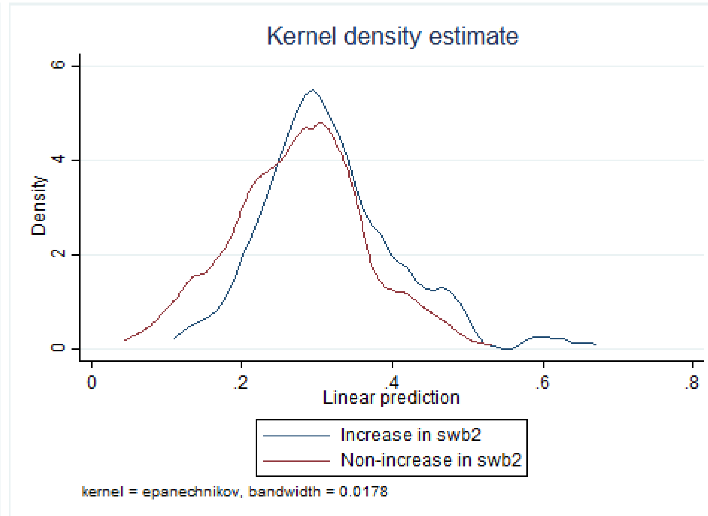


(c) Expansive Set

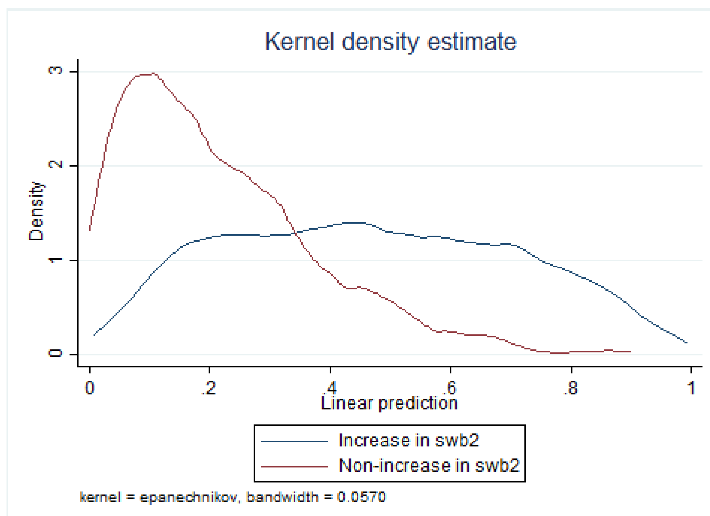
Figure 8 – Kernel Density Curve for SWB2 Estimation (Penalized Logistic Regression)



(a) Restricted Set

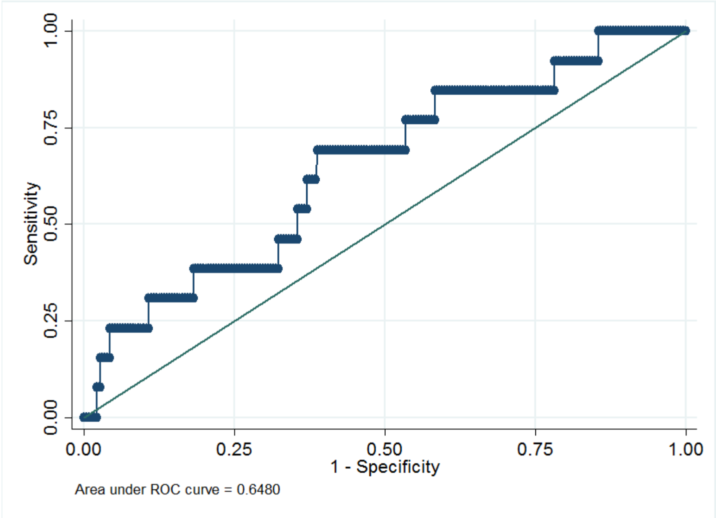


(b) Medium Set

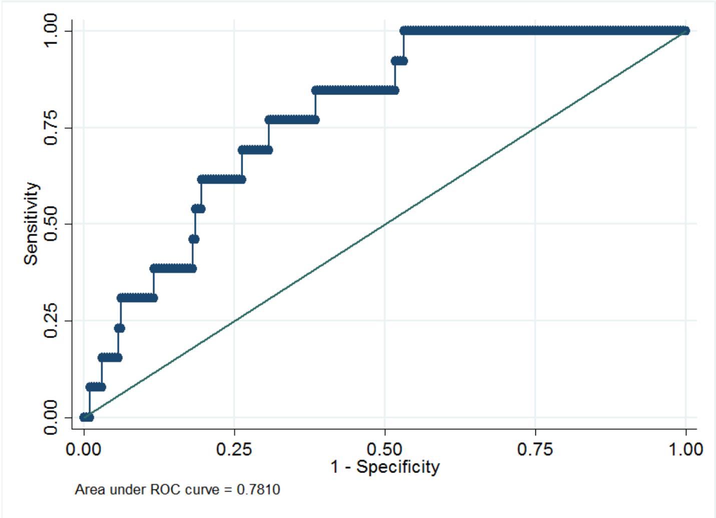


(c) Expansive Set

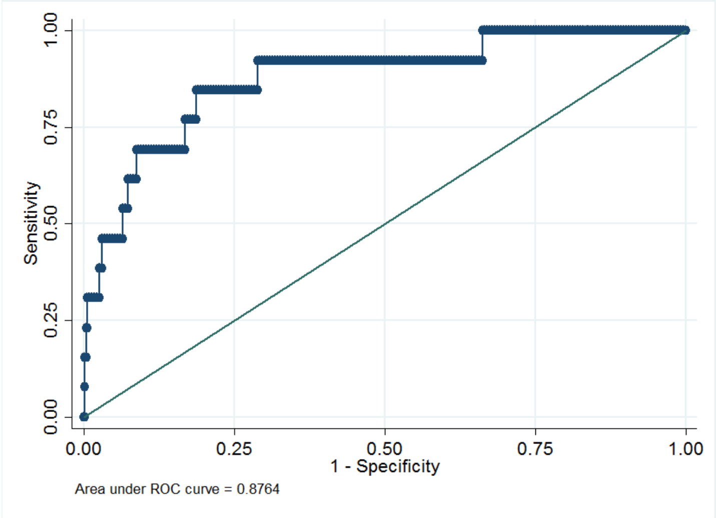
Figure 9 – ROC Curve for SWB1 Estimation (Logistic Regression)



(a) Restricted Set

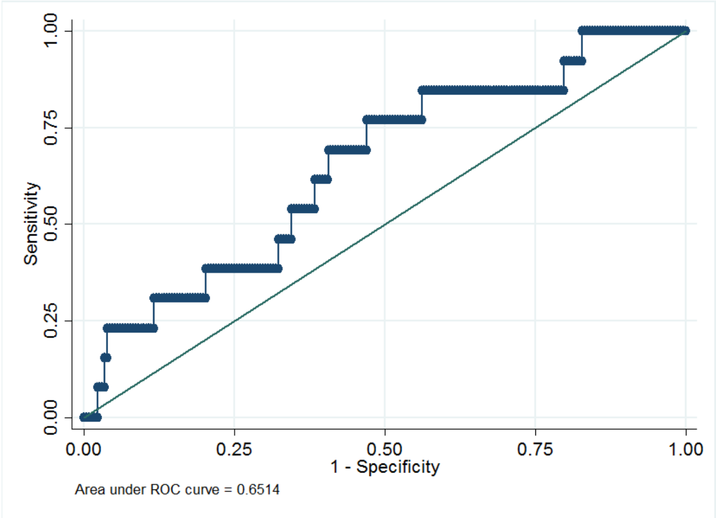


(b) Medium Set

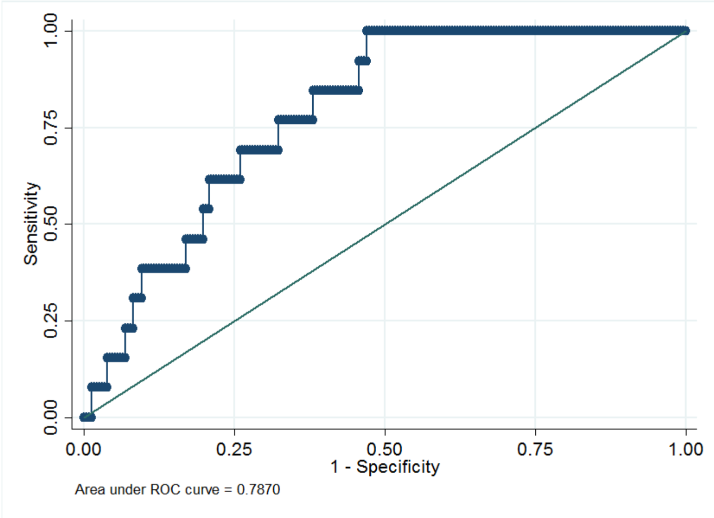


(c) Expansive Set

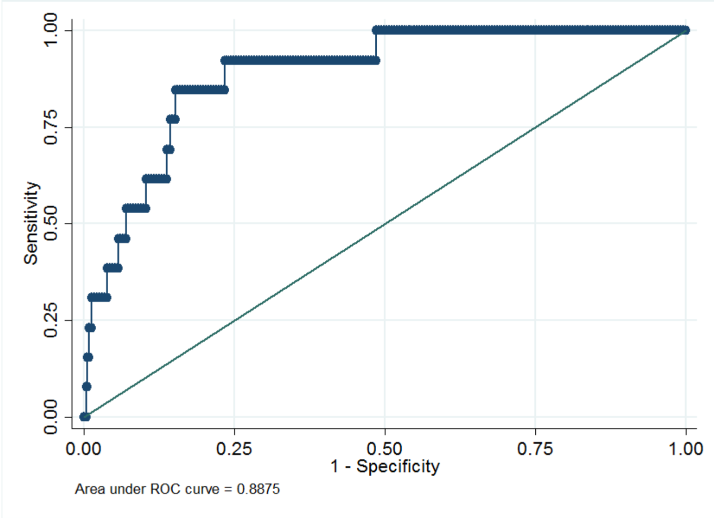
Figure 10 – ROC Curve for SWB1 Estimation (Probit Regression)



(a) Restricted Set

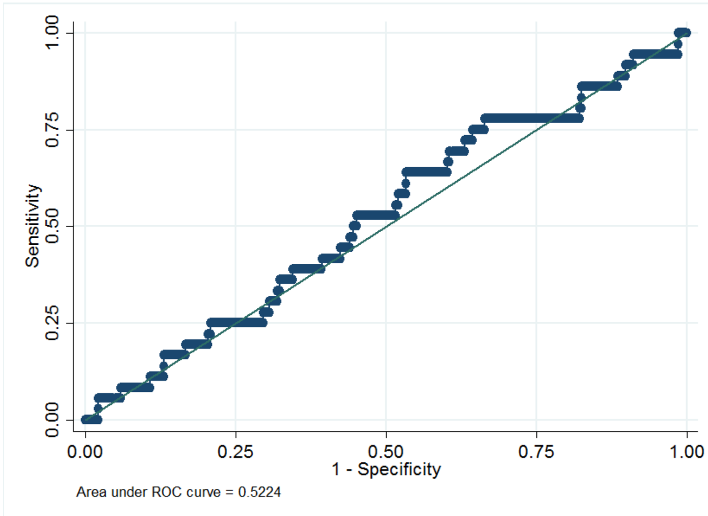


(b) Medium Set

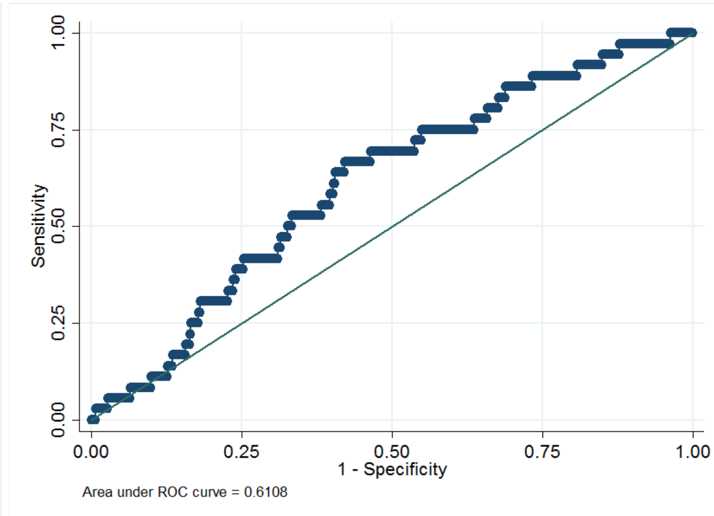


(c) Expansive Set

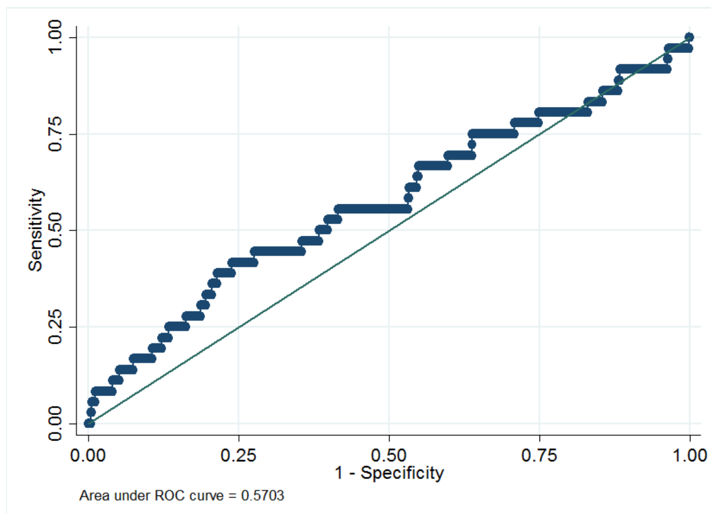
Figure 11 – ROC Curve for SWB1 Estimation (Penalized Logistic Regression)



(a) Restricted Set

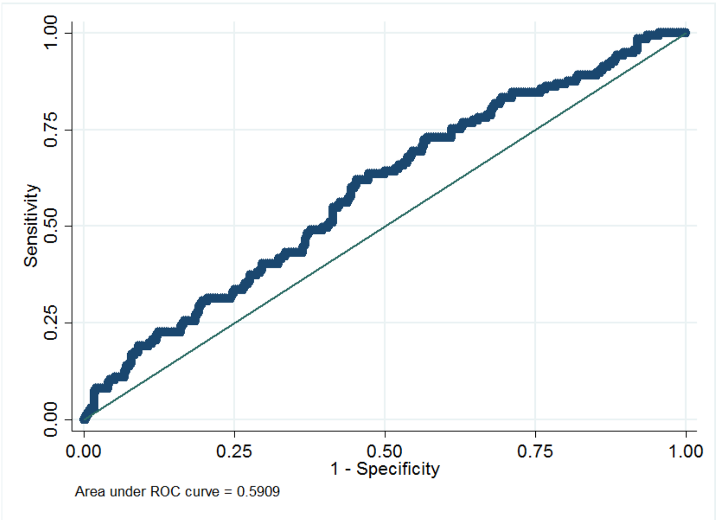


(b) Medium Set

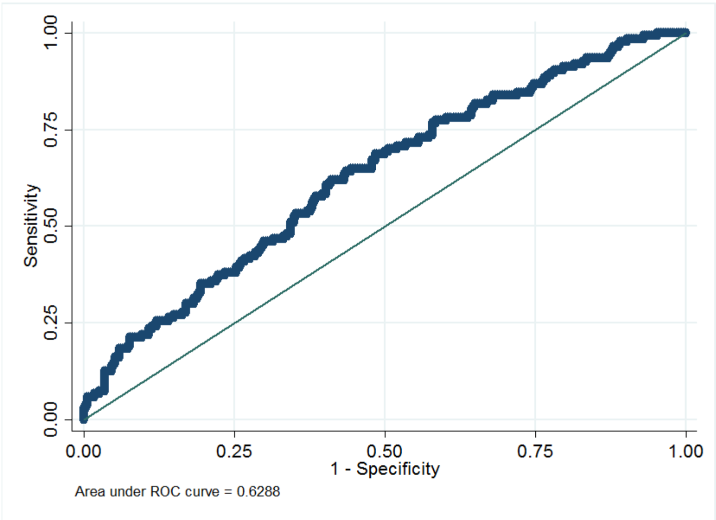


(c) Expansive Set

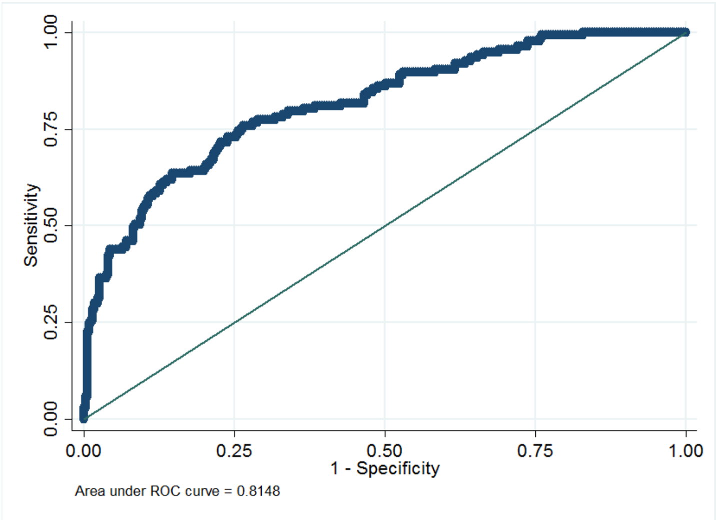
Figure 12 – ROC Curve for SWB2 Estimation (Logistic Regression)



(b) Restricted Set

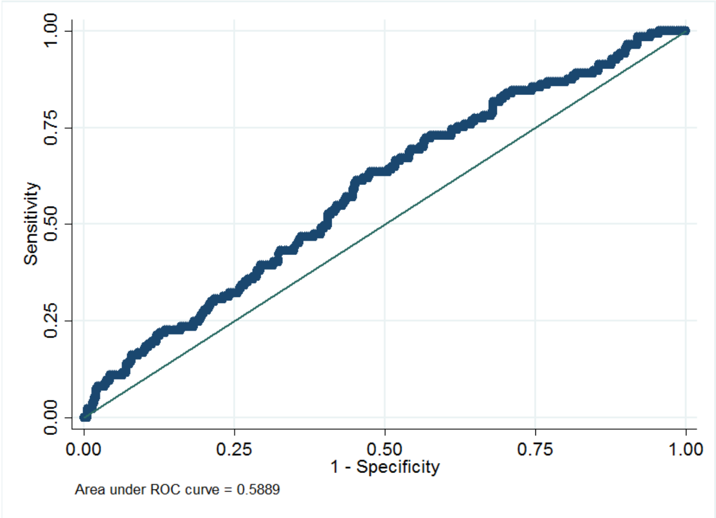


(b) Medium Set

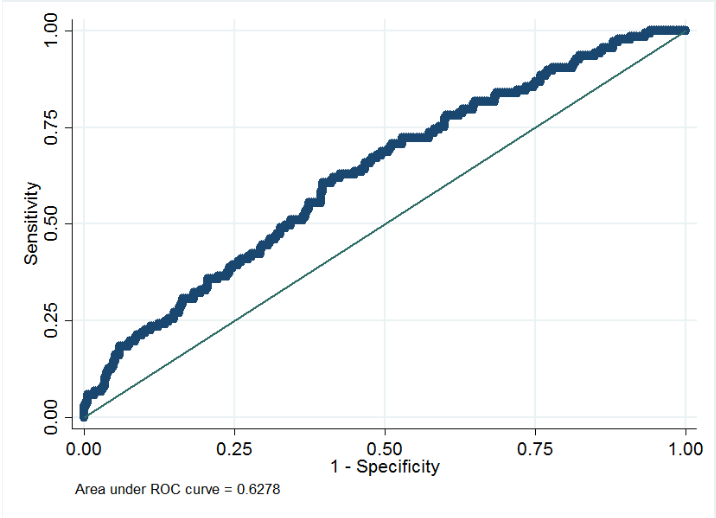


(c) Expansive Set

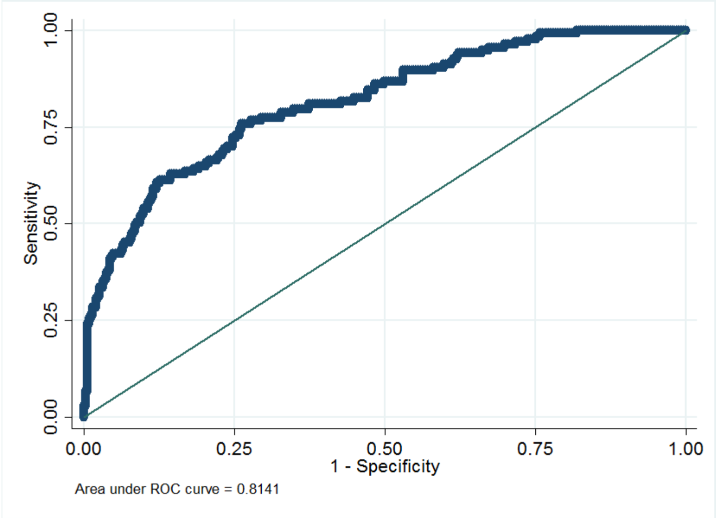
Figure 13 – ROC Curve for SWB2 Estimation (Probit Regression)



(b) Restricted Set

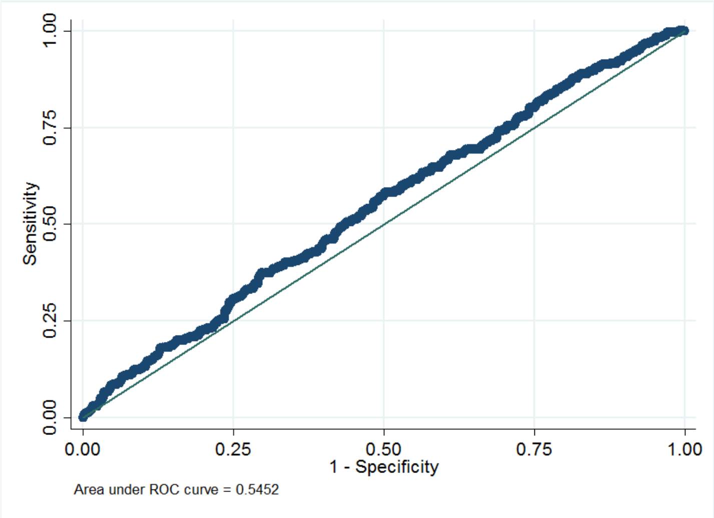


(b) Medium Set

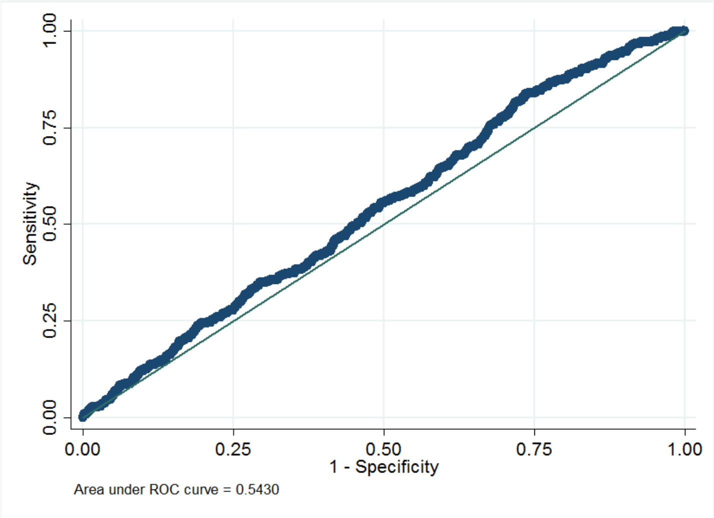


(c) Expansive Set

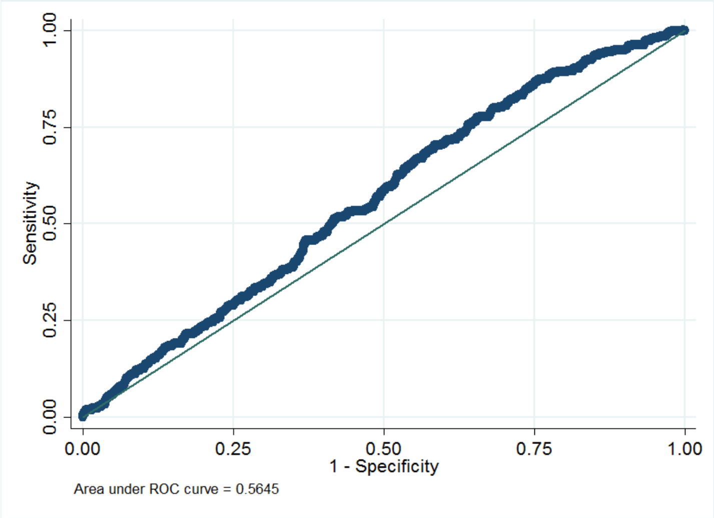
Figure 14 – ROC Curve for SWB2 Estimation (Penalized Logistic Regression)



(b) Restricted Set



(b) Medium Set



(c) Expansive Set