

Duong, Be Thanh; Sariyev, Orkhan; Zeller, Manfred

Article — Published Version

Can Weight of Evidence Method Improve Poverty Targeting?

Social Indicators Research

Provided in Cooperation with:

Springer Nature

Suggested Citation: Duong, Be Thanh; Sariyev, Orkhan; Zeller, Manfred (2025) : Can Weight of Evidence Method Improve Poverty Targeting?, Social Indicators Research, ISSN 1573-0921, Springer Netherlands, Dordrecht, Vol. 178, Iss. 1, pp. 305-333, <https://doi.org/10.1007/s11205-025-03576-z>

This Version is available at:

<https://hdl.handle.net/10419/323668>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<http://creativecommons.org/licenses/by/4.0/>



Can Weight of Evidence Method Improve Poverty Targeting?

Be Thanh Duong^{1,2} · Orkhan Sariyev¹ · Manfred Zeller¹

Accepted: 7 March 2025 / Published online: 6 April 2025
© The Author(s) 2025

Abstract

This study explores the potential of adapting a WOE (Logit) method, a combination of techniques such as weight of evidence, information value, logit regression, and a score-scaling approach based on doubling the odds, for poverty targeting. To verify the effectiveness and accuracy of this method, the study applies it to develop targeting tools based on international and Vietnamese national poverty lines, and compares their accuracy to commonly used poverty-targeting tools such as the Simple Poverty Scorecard, Proxy Means Test, and Poverty Assessment Tool. The study uses coverage rate as a criterion to compare the accuracy of targeting tools in identifying the poor at household and individual levels. The results indicate that targeting tools constructed using the WOE (Logit) method outperform those developed using other methods. Depending on the poverty line, this method improves poverty identification accuracy by 1.9–5.9 percentage points for households and 1.5–3.3 percentage points for individuals compared to other methods. Therefore, this study contributes an additional method for constructing targeting tools with high accuracy alongside existing approaches.

Keywords Poverty targeting tool · Proxy means test · Simple poverty scorecard · Poverty assessment tool · Credit risk scorecard · Poverty targeting

1 Introduction

To identify impoverished people eligible for certain poverty reduction programs to receive cash or in-kind transfers aimed at improving their well-being, many countries prefer low-cost but effective poverty assessment and targeting tools such as Proxy Means Test (PMT), Poverty Assessment Tool (PAT), or Simple Poverty Scorecard (Scorecard) (Etang et al., 2022; Schreiner, 2023; Zeller et al., 2006). These are questionnaires with specific answers

✉ Be Thanh Duong
dbthanh1991@gmail.com; duong.thanh@uni-hohenheim.de

Orkhan Sariyev
o.sariyev@uni-hohenheim.de

Manfred Zeller
manfred.zeller@uni-hohenheim.de

¹ Hans-Ruthenberg-Institute, University of Hohenheim, Stuttgart, Germany

² Faculty of Agriculture and Rural Development, Kien Giang University, Kien Giang, Vietnam

and related scores for each question, enabling authorities to quickly predict whether households are above or below a given poverty line through straightforward and verifiable questions. For brevity, in this study, we refer to these tools collectively as targeting tools. When using targeting tools for scoring, households with total scores below a certain benchmark are classified as poor, and vice versa. This transparency makes users and the public more likely to accept and trust the accuracy of targeting tools in identifying the poor (Attanasio et al., 2023; Leite & George, 2014; Schreiner et al., 2014).

To create transparent targeting tools, variables must be transformed into categorical variables, as each question in a targeting tool requires specific answers with assigned scores. Therefore, transforming original variables into categorical ones is a fundamental step in the construction process of targeting tools. Values within a variable can be grouped in various ways, resulting in numerous new categorical variables from a single original variable. Categorization for variables with wide-ranging values is even more challenging, as without a specific criterion, it is difficult to determine the best category groupings. Each new categorical variable can affect the predictability of targeting tools differently. Although the impact of a single variable may be minor, ineffective groupings across many variables can prevent targeting tools from achieving optimal predictability. Additionally, the way categories are grouped for a variable can influence statistical values of prediction models, such as c-statistic values, adjusted R-squared, or Akaike Information Criterion. Typically, the best indicators for targeting tools are selected through a stepwise process, in which the fractional change in these statistical values serves as one of the key selection criteria (Ohlenburg et al., 2022; Schreiner, 2023; Zeller et al., 2005). Consequently, how categories are grouped within a variable can also affect the indicator selection for targeting tools, ultimately affecting their predictability. However, these aspects remain underexplored, as most existing studies primarily focus on finding optimal prediction algorithms to enhance the accuracy of targeting tools. These studies often compare various predictive modeling techniques, ranging from traditional statistical methods such as linear probability models, logistic regression, ordinary least squares (OLS), and quantile regression (Linh & Baulch, 2011; Schreiner, 2015; Skoufias et al., 2020; Zeller et al., 2006) to more advanced machine learning algorithms (Alsharkawi et al., 2021; Kambuya, 2017; McBride & Nichols, 2015). Nevertheless, no traditional regression model consistently outperforms others across different contexts (Etang et al., 2022; Schreiner, 2015). Furthermore, the accuracy gains from using sophisticated machine learning algorithms over simpler regression techniques are often marginal (Hand, 2006).

Weight of Evidence (WOE) and Information Value (IV) techniques are often used to address these binning issues in the development of credit risk scorecards (Anderson, 2007; Siddiqi, 2016). The WOE technique helps assess the strength of the relationship between groups within a variable, known as bins, and a given dependent variable (Lund, 2016). The IV, on the other hand, is a useful measure for evaluating and comparing the predictive power of variables, and it can serve as a benchmark to select optimal category groupings for indicators (Siddiqi, 2006). Incorporating these techniques facilitates the variable transformation process, helping create monotonic variables with high predictive power. Moreover, applying the WOE technique in variable transformations has been shown to improve the performance of predictive modeling in credit scoring (Chen et al., 2020; Pomazanov, 2023; Sharma, 2011).

Despite the proven effectiveness of WOE and IV techniques in credit risk assessment, their application in poverty identification remains underexplored. Therefore, this study aims to examine the feasibility and efficacy of applying these techniques to poverty identification. In credit scoring, logistic regression and the score-scaling approach

of doubling the odds are often used along with WOE and IV techniques to construct scorecards that predict the default probabilities of customers (Anderson, 2007; Siddiqi, 2016). Hence, we apply a combination of these approaches for identifying the poor. For brevity, we refer to this combination of approaches as the WOE (Logit) method in this study, which includes WOE, IV, Logit, and the score-scaling approach of doubling the odds. Targeting tools developed using this method are called WOE tools. To verify the accuracy of WOE tools, we compare them to other commonly used targeting tools such as PMT, PAT, and Scorecard. For clarity and distinction, we refer to the methods used to construct these targeting tools as the PMT (OLS) method, PMT (Quantile) method, and Scorecard (Logit) method, respectively. The coverage rate, or poverty inclusion rate, is a commonly used criterion to define the accuracy and benchmark score of targeting tools (Etang et al., 2022; IRIS Center, 2005; Schreiner, 2023). Therefore, we apply this criterion to measure and compare the accuracy of targeting tools while holding the predicted poverty rate.

Based on an intensive literature review, to the best of our knowledge, this study is the first to apply the WOE (Logit) method to construct targeting tools. By doing so, it introduces another approach to poverty identification alongside existing methods. This method could help optimize input variables to develop high-accuracy targeting tools. In this study, we apply the four aforementioned methods to develop targeting tools based on two poverty lines: Vietnam's national poverty line and the international poverty line. Our focus is exclusively on building tools to identify the poor in the ethnic minorities, which have the highest poverty rate in Vietnam. Although ethnic minorities¹ comprise only 14% of Vietnam's population, they constitute half of the country's poor. Thus, this group is of special concern to the Vietnamese government (CEMA, 2020; CEMA et al., 2015). The accuracy of targeting tools will be compared not only at the household level but also at the individual level, as effective targeting tools should target households with more poor members rather than those with fewer members.

The remaining sections of the paper are organized as follows: Sect. 2 presents the data source and methodology; Sect. 3 presents the study's results and discussion; and Sect. 4 summarizes the major findings.

2 Data and Methodology

2.1 Data

This study uses data from Vietnam Household Living Standard Survey in 2018 (VHLSS 2018), conducted by the General Statistics Office of Vietnam. This official dataset is useful for developing targeting tools because it includes necessary information such as household demographics, homestead characteristics (e.g., type of house or toilet), durable goods, education, employment, and income (General Statistics Office, 2019). The primary focus of this study is the ethnic minority community, which has the highest poverty rate in Vietnam. Therefore, we analyze data from 7,863 ethnic minority households.

The VHLSS 2018 dataset contains over 600 indicators covering various aspects of households, including demographics, housing, education, employment, and assets.

¹ Vietnam has 53 ethnic minorities and one ethnic majority (Kinh).

First, we remove indicators with more than 10% missing values.² We then select 129 straightforward and verifiable indicators that exhibit a strong correlation with household income, expenditure, or poverty status. After identifying candidate variables, we split the dataset using stratified random sampling into a calibration sample (70%) and a validation sample (30%). We use the calibration sample to construct targeting tools, and the validation sample to evaluate their accuracy.

2.2 Poverty line

The study evaluates the accuracy of four methods for constructing targeting tools based on national and international poverty lines. However, the Vietnamese government has different national poverty lines for rural and urban areas. Since 90% of the ethnic minority population and 95% of poor ethnic minorities reside in rural areas (CEMA, 2020; General Statistics Office, 2019), this study uses only the rural poverty line for all ethnic minority communities (Table 1).

2.3 Methodology

2.3.1 General Steps in Developing Poverty Targeting Tools

To verify the feasibility and effectiveness of the WOE (Logit) method, we compare it with three widely used methods: Scorecard (Logit), PMT (OLS), and PMT (Quantile). Each method is used to develop two distinct targeting tools based on national and international poverty lines. To ensure a fair comparison across all methods, we limit the number of indicators in each targeting tool to ten. Table 2 presents a general step-by-step framework for constructing targeting tools.

Step 1: Grouping responses to indicators

Each question in targeting tools must have a specific answer with a designated score. This requirement ensures transparency and helps enumerators accurately identify the poor in the field. To meet this requirement, input variables for constructing targeting tools should be transformed to categorical variables. Numeric variables can be transformed into many new categorical variables, as there are various ways to group categories. Moreover, creating categories for variables with a wide range of values is challenging, as there are countless ways to form category groupings. However, existing studies have not provided detailed guidance on how to effectively create these categorical variables. Therefore, we generate new categories for variables based on those in the existing tool of the Vietnamese government. To maintain practicality, we limit the number of categories for each variable to seven, aligning with the Vietnamese government's tool. Next, we conduct a cross-tabulation analysis between these new categorical variables and a given poverty status to assess the relationship between the categories and poverty. If the categories do not align with the given poverty status, we regroup them accordingly to ensure they make sense and represent at least 5% of the sample.

For example, in Table 3, we can categorize household size into six groups: "1 member", "2 members", "3 members", "4 members", "5 members", "6 members", and "7 members or more". By creating a cross-tabulation between this variable and the national poverty

² Regarding variables include missing values with less than 10%, we replace their missing values with their mean or mode values.

Table 1 Poverty rates of the Vietnamese ethnic minority community in terms of income in 2018. *Source:* Authors' computations based on the VHLSS 2018 data

Household poverty status ^a based on income and expenditure		Poor	Non-poor	All
National poverty line (Income/person/day \leq 2.92 USD in purchasing power parity) ^b	N	1,293 (16.4%)	6,570 (83.6%)	7,863 (100%)
International poverty line (Expenditure/person/day \leq 1.9 USD in purchasing power parity)	N	1,054 (13.4%)	6,809 (86.6%)	7,863 (100%)

^aThe poverty status is defined based on income or expenditure, so the poverty rate in Table 1 is different from the poverty rate published by the Vietnamese government, which is based on a multi-dimensional poverty status

^bThis is based on the PPP conversion factor of 7,891.2 from the World Bank in 2018 to convert Vietnam Dong (VND) to USD

Table 2 Steps for constructing poverty targeting tools by the method

Step	Activity	Scorecard (Logit) method	PMT (OLS) method	PMT (Quantile) method	WOE (Logit) method
1	Group responses to indicators	Manual categorization based on cross-tabulation			WOE & IV technique
2	Indicator selection	“MaxC” stepwise	“MaxR” stepwise	PMT’s indicators	“MaxC” stepwise
3	Regression approach	Logit	OLS	Quantile	Logit
4	Generate scorecards	Scorecard (Logit) method	PMT (OLS) method	PMT (OLS) method	WOE (Logit) method
5	Estimate accuracy	Coverage rate in identifying poor households and people based on national and international poverty lines			
	Source for more details of each method	Schreiner (2023)	Nguyen & Tran (2018)	Zeller et al (2005)	Siddiqi (2016)

Table 3 An example of manual categorization by using cross-tabulation analysis. *Source:* Authors' computations based on the VHLSS 2018 data

<i>Initial categorization</i>							
Category (Number of household members)	1	2	3	4	5	6	≥ 7
Number of poor and non-poor (Poor: Non-poor)	28:145	58:548	78:893	215:1298	201:810	146:524	180:381
Poor rate of each category (%)	16.2	9.6	8.0	14.2	19.9	21.8	32.1
<i>Second categorization</i>							
Category (Number of household members)		<3	3	4	5	6	≥ 7
Poor rate of each category (%)		11.0	8.0	14.2	19.9	21.8	32.1
<i>Final categorization</i>							
Category (Number of household members)			<4	4	5	6	≥ 7
Poor rate of each category (%)			9.4	14.2	19.9	21.8	32.1

status, we find the following poverty rates for each category: 16.2, 9.6, 8, 14.2, 19.9, 21.8, and 32.1 percentage points, respectively. Since the first category accounts for only 3.1% of the population, we merge it with the second category to form a new category: “<3 members”. This new category has a poverty rate of 11 percentage points. However, the poverty rates of these categories are inconsistent and contradict common sense, as the category “<3 members” has a higher poverty rate than the category “3 members”. To address this, we merge the new category “<3 members” with the third category “3 members” resulting in the category “<4 members” with a poverty rate of 9.4 percentage points. Ultimately, the household size variable is categorized into five groups: “<4 members”, “4 members”, “5 members”, “6 members”, and “ ≥ 7 members”. The poverty rates for these categories are now consistent and make sense: 9.4, 14.2, 19.9, 21.8, and 32.1 percentage points.

Step 2: Selecting indicators for the prediction model

After obtaining candidate variables in Step 1, this step selects the ten best indicators for prediction models. Depending on the methodology, these indicators can be selected using the “MaxC” or “MaxR” stepwise approaches. The “MaxC” approach refers to a forward stepwise method performed manually to select indicators based on both statistical and non-statistical criteria. The statistical criterion is the c-statistic, also known as the concentration index (Ravallion, 2007) or the area under the receiver operating characteristic curve (Wodon, 1997). A higher value of this criterion indicates better model performance (Baulch, 2002). Non-statistical criteria are our judgment about various aspects of the indicators, such as user acceptability, simplicity, verifiability, applicability, variety, and so on. For more details about the “MaxC” forward stepwise approach, please refer to Schreiner (2023). The “MaxR” approach is similar to “MaxC” approach, but it uses adjusted R-squared as the statistical criterion. For more details about the “MaxR” approach, please refer to Zeller et al. (2005).

Step 3: Training prediction model

In this step, we use selected indicators to develop prediction models and obtain weights for those indicators. Depending on the methodology, the prediction models may use

Logit, OLS, or Quantile regression algorithms. Detailed information about the regression algorithms and dependent variables used will be provided in the description of each method below.

Step 4: Transforming the prediction model's coefficients to score

Based on the results of prediction models, the coefficients of indicators can be transformed into scores to form targeting tools in three simple steps. First, we switch reference categories of variables if their coefficients are negative. This ensures all coefficients are positive. Second, we multiply the coefficients of the variables by 100. Finally, we round the scores obtained. The sum of these scores (excluding the intercept's coefficient) is the total score for predicting the poverty status of households. This common and straightforward approach is widely used to create targeting tools and is also applied by the Vietnamese government for their existing tools (Nguyen & Lo, 2016). The Scorecard and WOE (Logit) methods use different approaches to transform coefficients into scores, which will be explained later.

Step 5: Estimating the tool's accuracy

We use coverage rate as a key criterion to compare the accuracy of targeting tools. The coverage rate, also known as the inclusion of the poor or poverty accuracy, is currently used by the Vietnamese government to assess the accuracy of their targeting tools (IRIS Center, 2005; Nguyen & Lo, 2016). The coverage rate is the percentage of households correctly predicted as poor by a given targeting tool, expressed as a percentage of the total number of observed poor households. In addition, we also measure the coverage rate of targeting tools at the person level for comparison. This means we also calculate the coverage rate of targeting tools when identifying poor individuals. This is important because identifying poor households with larger family sizes is more beneficial than identifying those with fewer members.

We measure the coverage rate of targeting tools when they perform in validation samples. This helps us avoid the overfitting issue because targeting tools certainly perform well in calibration samples, which are used to construct them. It also allows us to gauge the accuracy of targeting tools in identifying the poor within an unobserved population. Additionally, we apply a bootstrap technique to estimate the standard error of the coverage and targeted rates by calculating those measures and their differences for the 1,000 repetitions from the bootstrap sample (Chen & Schreiner, 2009; Efron & Tibshirani, 1994). This approach provides a reliable assessment of tools' performance.

For a fair comparison, we compare the coverage rates at a cut-off with the same predicted poverty rate, which is the rate of the targeted population in validation samples. Due to the granularity of the scores, it is challenging to obtain cut-off scores with the same predicted poverty rate across all targeting tools. Therefore, we fix the poverty rate predicted by a specific targeting tool and use linear interpolation to estimate the coverage rates of the remaining targeting tools.

2.3.2 The Scorecard (Logit) Method

Scorecards are also developed based on the principles of the PMT (OLS) method. Scorecard (Logit) method uses the scores of proxy indicators to construct Scorecards for identifying the poor. Schreiner (2010) applies logistic regression to estimate the weights of proxy indicators, and uses his own method to convert the indicators' weights to non-negative

integer scores. Scorecards are straightforward as they use only 10 indicators to ascertain the poor, with a maximum total score of 100. This allows for a quick and easy calculation of households' total scores in the field. Scorecards have been developed for over 60 countries (Chen & Schreiner, 2009; Schreiner, 2006, 2023). Scorecards are still being updated and many countries continue to apply this tool. Therefore, it would be an omission not to compare the WOE (Logit) method to the Scorecard (Logit) method. For a fair comparison, we try to replicate Schreiner's method (2023) to construct Scorecards by following the below steps.

Step 1: We proceed analog to Step 1 in Sect. 2.3.1.

Step 2: Ten indicators of both these models are selected by applying the "MaxC" forward stepwise approach as mentioned in Step 2 in Sect. 2.3.1.

Step 3: We use logistic models to predict the probability of households falling below national and international poverty lines for constructing national and international prediction models respectively.

Step 4: The Scorecard (Logit) method generates national and international Scorecards by converting coefficients of their prediction models into non-negative integer scores, so that lower total scores correspond to higher poverty probabilities and vice versa. The maximum total score in scorecards is 100. Many studies have shown that this approach produces accurate scorecards for targeting the poor (Chen & Schreiner, 2009; Schreiner, 2012, 2023). For more details on how the Scorecard (Logit) method transforms coefficients into scores, please refer to Schreiner's approach (2022).

Step 5: We proceed analog to Step 5 in Sect. 2.3.1.

2.3.3 The PMT (OLS) Method

The PMT is a fundamental and age-old tool in the field of poverty targeting. The basic PMT is constructed using an OLS algorithm to predict households' income or expenditure. The weights of indicators estimated by OLS models are transformed into proxy scores to predict the poor at given poverty lines. The PMT emerged as an effective tool in social welfare programs in Latin America due to its effectiveness and accuracy (Grosh et al., 1995). It is widely applied in many countries (e.g., Burchi et al., 2018; Maité, 2012; Nguyen & Lo, 2016; Zeller et al., 2005; Rawlings & Rubio, 2003). The PMT (OLS) method is also commonly used to develop targeting tools for Vietnam and is currently used by the Vietnamese government (Nguyen & Lo, 2016; Nguyen & Tran, 2018; USAID, 2006). We apply the PMT (OLS) method to develop PMTs in this study by following these steps:

Step 1: We proceed analog to Step 1 in Sect. 2.3.1

Step 2: To obtain the best ten variables for prediction models, the PMT (OLS) method applies a "MaxR" forward stepwise approach which is described in Step 2 of Sect. 2.3.1.

Step 3: The PMT (OLS) method uses OLS regressions for predictive modeling and to estimate the weights of indicators. In the national model, which is based on the national poverty line, the dependent variable is the natural logarithm of income, while the international model uses the natural logarithm of expenditure.

Steps 4 and 5: We proceed analog to Steps 4 and 5, respectively, in Sect. 2.3.1.

2.3.4 The PMT (Quantile) Method

PATs are other versions of PMTs as they are also developed based on PMT principles. PATs for Vietnam and about 28 other countries have been developed by the Center for Institutional Reform and the Informal Sector (IRIS) at the University of Maryland. After applying the BPAC (Balanced Poverty Accuracy Criterion) to judge the accuracy of various methods, IRIS (2005) recommends the quantile regression as a higher accuracy algorithm for building PMTs compared to OLS, linear probability, and probit regressions (Zeller et al., 2006). Therefore, this study refers to the PMT method using the quantile regression as the PMT (Quantile) method. Below are the steps to construct PATs in this study.

Step 1: We proceed analog to Step 1 in Sect. 2.3.1

Step 2: The forward stepwise approach is not suitable to select optimal indicators for quantile regressions. Therefore, we use the same indicators selected in the PMT (OLS) method.

Step 3: The PMT (Quantile) method uses the same independent and dependent variables as the PMT (OLS) method, but it applies quantile regressions. To select the best percentiles for quantile regressions, we use the *c*-statistic, as mentioned in Step 2 of Sect. 2.3.1. This means that the selected percentiles are those that yield the highest *c*-statistic values for the quantile regressions.

Steps 4 and 5: We proceed analog to Steps 4 and 5, respectively, in Sect. 2.3.1.

2.3.5 The WOE (Logit) Method

In the credit scoring field, there are various approaches to constructing credit scorecards. The common one is the combination of various techniques such as WOE, IV, logistic regression, and the score-scaling approach based on doubling the odds. In this study, we refer to the method of combining all these techniques as the WOE (Logit) method, and targeting tools built by this method are called WOE tools. Below are the steps of the WOE (Logit) method.

Step 1: This step is different from other methods in this study, as the WOE (Logit) method does not use categorical variables and applies WOE and IV techniques for transforming and screening out variables.

The WOE technique is a variable transformation technique based on a concept from information theory (Kullback, 1997). It converts a variable into a more informative and monotonic variable for constructing prediction models by grouping the values of that

variable into bins (or groups) and assigning WOE values to each bin. In credit scoring, this technique measures the separation of good and bad customers (Anderson, 2007). In this study, we use the WOE technique to assess the strength of the relationship between a bin and poverty status. Bins with positive WOE values indicate a stronger association with poor status, while negative values suggest a stronger association with non-poor status. A higher positive or lower negative WOE magnitude implies better identification of the respective group. If the differentiation between the WOE values of bins in a variable is high, the predictive power (IV) of that variable will also be high (Lund, 2016). Therefore, during the binning process, we aim to maximize the differentiation between the WOE values of bins within a variable. This serves as a primary benchmark for creating optimal bins, as it helps the variable achieve its highest predictive power. However, other criteria, such as monotonicity and parsimony, will also be considered in subsequent steps to ensure their practical applicability. In this study, a WOE value of a bin is defined as follows:

$$WOE_i = Ln \left[\left(\frac{N_i}{\sum N} \right) \div \left(\frac{P_i}{\sum P} \right) \right] = Ln \left(\frac{\text{Poor distribution}_i}{\text{Non - poor distribution}_i} \right), \quad (1)$$

where P=quantity of poor; N=quantity of non-poor; and i=index of the bin. The precondition is that both N_i and P_i must be non-zero values.

In this research, the IV of an indicator represents the strength of that indicator in predicting the likelihood of households being poor (Kullback, 1997). Therefore, the IV is a useful index to eliminate indicators with weak predictive power, which are indicators with an IV below 0.01 (Siddiqi, 2006). This reduces the number of input variables in the stepwise process. The indicator's IV is the sum of the IVs of all bins in that indicator. We compute the IV of a bin based on the calculated WOE values. A variable with a higher IV indicates greater predictive power. Since different bin combinations produce varying IVs, we adjust the bins to maximize the IV for each variable while ensuring compliance with the criteria outlined in the binning process below.

The IV of a bin is defined as (Anderson, 2007, p. 193):

$$IV = \sum_{i=1}^n \left(\frac{P_i}{\sum P} - \frac{N_i}{\sum N} \right) \times WOE_i \quad (2)$$

The steps³ for binning a variable and calculating its WOE and IV values in this study are adapted as follows:

- i. Determine class variables: categorical or continuous. If variables are continuous, divide the data of variables into ten or fewer bins based on the distribution, making sure that each bin has at least 5% of the cases. If variables are categorical, move to Step ii.
- ii. Find the number of poor and non-poor values in each group.
- iii. Calculate poor and non-poor distributions and use them to calculate WOE values by applying Eq. (1).
- iv. Adjust the bins based on these criteria: (1) Distinct WOE values: Combine adjacent bins with similar WOE values to ensure that the WOE values are distinct for each bin;

³ Actually, instead of manually performing Steps i to iii to get bins, we use a technique in Machine Learning namely Decision Tree. This technique can capture optimal break points for creating ideal candidate bins. In the next steps, these candidate bins are manually refined based on criteria in Step iv to become final bins, and calculate their IV.

- (2) Monotonicity: Ensure that the WOE values of bins within a variable follow either an ascending or descending order. If not, revise the way adjacent bins are combined;
- (3) Predictive power: Maximize the differentiation between the WOE values of bins within a variable to achieve high-predictive power variables;
- (4) Parsimony and practicability: Consider the ease and user acceptability of these bins in practice when modifying them. For example, variables with bins that have odd values or many bins can achieve a high IV, but it may be difficult to collect or verify this information in practice.
- v. Compute the IVs of bins in variables based on Eq. (2).
 - vi. To calculate the IV of a variable, add up the IVs of all the bins in that variable.

Step 2: The best indicators for national and international prediction models are selected by applying the “MaxC” forward stepwise process outlined in Step 2 of Sect. 2.3.1.

Step 3: After applying WOE and IV techniques, we obtain new variables with WOE values, called WOE variables. Instead of using categorical variables as in other methods, we use these WOE variables for logistic models to predict the probability of households falling below the national and international poverty lines.

Step 4: After obtaining the national and international prediction models, we use a score-scaling approach of doubling the odds, a commonly applied method in the credit scoring domain, to develop national and international WOE tools, respectively. Details about the mathematical transformation of this approach can be found in (Siddiqi, 2016, p. 241). Here, we briefly present the equation transformed from this approach, Eq. (3), to calculate the score for each response to indicators.

$$Score_{ij} = \left(WOE_j \times \beta_i + \frac{\alpha}{n} \right) \times Factor + \frac{Offset}{n} \quad (3)$$

where: α is the intercept of the logistic regression;

β is the regression coefficient of the indicator;

n is the number of indicators in the scorecard;

i is the index of indicator from 1 to n ;

k is the maximum number of bins in an i indicator;

j is the index of bins in i indicators;

Offset = Target Score – [Factor \times ln (Target Odds)];

Factor = PDO \div ln (2);

and WOE is the weight of evidence of j bin in an i indicator.

Target Score and Target Odds are the desired total score and desired poverty odds selected at the outset to scale up scores for WOE tools. Target Score, Target Odds, and PDOs are self-selected values. Changing these values only affects the magnitude of the total scores and does not significantly impact the accuracy performance (Siddiqi, 2016). A PDO of 20 is a common value used in the credit scoring field. However, credit scorecards aim to target non-default customers who have large total scores, while WOE tools aim to identify the poor. For more intuitive results, we select a negative PDO, -20 , so that households with small total scores will be predicted as poor, and vice versa.

We choose a Target Score of 300 instead of the usual 500 or 600 because our WOE tools consist of 10 indicators. A Target Score of 300 allows for the lowest possible total

scores while avoiding any negative scores in WOE tools. We use a target odds of 1 to simplify the determination of the optimal cut-off score. With a target odd of 1, the optimal cut-off score will fall within the range from the Target Score minus PDO to the Target Score plus PDO (e.g., 320–280). Also, selecting a target odds of 1 is for ease of explanation. For example, our selected Target Score and Target Odds are 300 and 1, respectively. This means that households with a predicted poverty odds of 1 (poor probability of 0.5) will receive a total score of 300. The point to double this odds is –20. Then, households with predicted poverty odds of 2 (poor probability of 0.667) have total scores of 280 (300 plus –20). Conversely, those with predicted poverty odds of 0.5 (poor probability of 0.333) have total scores of 320 (300 minus –20).

Step 5: We proceed analog to Step 5 in Sect. 2.3.1.

3 Result and Discussion

3.1 The Methods' Accuracy in Identifying Poor Households

Table 4 presents the coverage rate and predicted poverty rate of a national Scorecard constructed using the Scorecard (Logit) method when the cut-off score is set at 24. At this cut-off score, the national Scorecard achieves a coverage rate of 58.0 percentage points and a predicted poverty rate of 16.4 percentage points. This means that the national Scorecard identifies 16.4% of the population in the validation sample as poor households, of which 58.0% are actual poor households. Since the predicted poverty rate of the national Scorecard is the same as the observed national poverty line (16.4 percentage points), we hold the predicted poverty rate at 16.4 percentage points for comparing the coverage rates of all four national targeting tools.

For the remaining national targeting tools, since we were unable to find a cut-off score that achieves a predicted poverty rate of 16.4 percentage points, we applied the linear interpolation method to estimate their coverage rates. As a result, when targeting 16.4% of the population as poor households, the coverage rates for the national PMT, PAT, and WOE tools are 54.0, 55.6, and 59.9 percentage points, respectively, as shown in Table 4. These results indicate that the national WOE tool achieves the highest coverage rate in identifying poor households based on the national poverty line. Specifically, the coverage rate

Table 4 The accuracy of targeting tools in identifying poor households by the poverty line

	Scorecard	PMT	PAT	WOE tool
<i>National poverty line (in household)</i>				
Coverage rate	58.0	54.0	55.6	59.9
Predicted poverty rate	16.4	16.4	16.4	16.4
<i>International poverty line (in household)</i>				
Coverage rate	51.6	51.3	50.9	54.3
Predicted poverty rate	13.4	13.4	13.4	13.4

Unit: (%)

Table 5 The standard error and p-value of targeting tools in identifying poor households by the poverty line

National poverty line (in household)		Unit (%)	Std. error	p-value	International poverty line (in household)		Unit (%)	Std. error	p-value
Scorecard	Coverage rate	58.0	2.510	0.001	Scorecard	Coverage rate	51.6	2.719	0.001
	Predicted poverty rate	16.4	0.796	0.001		Predicted poverty rate	13.4	0.702	0.001
PMT	Coverage rate	53.9	2.504	0.001	PMT	Coverage rate	50.6	2.690	0.001
	Predicted poverty rate	16.2	0.785	0.001		Predicted poverty rate	13.1	0.663	0.001
PAT	Coverage rate	55.7	2.488	0.001	PAT	Coverage rate	50.6	2.704	0.001
	Predicted poverty rate	16.5	0.789	0.001		Predicted poverty rate	13.4	0.670	0.001
WOE tool	Coverage rate	60.6	2.529	0.001	WOE tool	Coverage rate	54.4	2.707	0.001
	Predicted poverty rate	16.6	0.816	0.001		Predicted poverty rate	13.5	0.697	0.001

of the national WOE tool is 1.9, 5.9, and 4.3 percentage points higher than those of the national Scorecard, PMT, and PAT, respectively.

Regarding the international poverty line, Table 4 shows that the international WOE tool achieves the highest accuracy among international targeting tools, with a coverage rate of 54.3 percentage points. In comparison, the international Scorecard and PMT have lower accuracy, with coverage rates of 51.6 and 51.3 percentage points, respectively, while the international PAT has the lowest accuracy, with a coverage rate of 50.9 percentage points. This means that the coverage rate of the international WOE tool is 2.7, 3.0, and 3.4 percentage points higher than those of the international Scorecard, PMT, and PAT, respectively.

To estimate standard errors for the coverage rates and predicted poverty rates of national and international targeting tools, the bootstrap sampling method requires the actual cut-off scores of these tools. However, as mentioned earlier, national targeting tools do not have exact cut-off scores that achieve the same predicted poverty rate, and neither do international targeting tools. Therefore, to estimate standard errors, we use cut-off scores that yield predicted poverty rates closest to the observed poverty rates. The results in Table 5 show that the standard errors of the national WOE tool differ only slightly from those of other national targeting tools, and its p-value is identical to those of the other national tools. Similarly, for the international poverty line, Table 5 indicates that the standard errors of the international WOE tool are comparable to those of other international targeting tools, and the p-values of all four international targeting tools are the same.

3.2 The Methods' Accuracy in Identifying Poor People

The accuracy comparison of the four methods in identifying poor individuals follows the same approach used for identifying poor households. We apply the previously constructed targeting tools to the corresponding poverty lines and calculate the coverage rate for each tool. The key difference is that the national and international poverty lines are now measured at the individual level (household member). As a result, these poverty lines are higher than those used in the household-level analysis.

Table 6 presents the coverage and predicted poverty rates of the national Scorecard, which are 64.1 and 20.8 percentage points, respectively. This means that when targeting 20.8% of the population at the individual level, the national Scorecard correctly identifies 64.1% of them as poor people. For the same targeted population share, the coverage rates of the national PMT, PAT, and WOE tools are 62.2, 62.7, and 65.5 percentage

Table 6 The accuracy of targeting tools in identifying poor people by the poverty line

	Scorecard	PMT	PAT	WOE tool
<i>National poverty line (in person)</i>				
Coverage rate	64.1	62.2	62.7	65.5
Predicted poverty rate	20.8	20.8	20.8	20.8
<i>International poverty line (in person)</i>				
Coverage rate	55.5	56.9	56.2	58.8
Predicted poverty rate	17.2	17.2	17.2	17.2

Unit: (%)

Table 7 The standard error and p-value of targeting tools in identifying poor people by the poverty line

National poverty line (in person)		Unit (%)	Std. error	p-value	International poverty line (in person)		Unit (%)	Std. error	p-value
Scorecard	Coverage rate	64.1	1.078	0.001	Scorecard	Coverage rate	55.5	1.173	0.001
	Predicted poverty rate	20.8	0.408	0.001		Predicted poverty rate	17.2	0.354	0.001
PMT	Coverage rate	61.2	1.119	0.001	PMT	Coverage rate	56.7	1.200	0.001
	Predicted poverty rate	20.3	0.407	0.001		Predicted poverty rate	17.0	0.356	0.001
PAT	Coverage rate	61.2	1.108	0.001	PAT	Coverage rate	55.1	1.194	0.001
	Predicted poverty rate	20.1	0.407	0.001		Predicted poverty rate	16.9	0.359	0.001
WOE tool	Coverage rate	64.2	1.082	0.001	WOE tool	Coverage rate	58.5	1.168	0.001
	Predicted poverty rate	20.2	0.391	0.001		Predicted poverty rate	17.0	0.352	0.001

points, respectively, as shown in Table 6. These results indicate that the national WOE tool achieves the highest accuracy in identifying poor individuals based on the national poverty line. Specifically, the coverage rate of the national WOE tool is 1.5, 3.3, and 2.8 percentage points higher than those of the national Scorecard, PMT, and PAT, respectively.

Regarding the identification of poor individuals based on the international poverty line, the international WOE tool also demonstrates the highest accuracy among all targeting tools. Specifically, its coverage rate exceeds those of the Scorecard, PMT, and PAT by 3.2, 1.9, and 2.6 percentage points, respectively.

In identifying poor individuals, both national and international WOE tools exhibit only minor differences in standard errors compared to those of other targeting tools, as shown in Table 7. Additionally, the p-values of the WOE tools are the same as those of the other targeting tools.

3.3 Discussion

The comparisons above show that the WOE (Logit) method outperforms the other three methods. However, due to differences among the four methods and the lack of controlled conditions, pinpointing the exact reasons for this outperformance is challenging. Nonetheless, we will highlight key differences between the methods to examine how these variations potentially influence accuracy and indicator selection in targeting tools.

The first difference lies in how responses for each indicator are binned. Other methods create categories based on the results of cross-tabulation, aiming to create groups that consistently align with poverty status. In the WOE (Logit) method, binning with the WOE technique does not stop at that requirement, but aims to maximize IV, the predictive power of an indicator. This approach often results in different binning patterns for variables with a wide range of values, such as “electricity bill,” “number of poultry,” “dependent ratio,” and so on. This is because without the criterion of IV, it is difficult to determine the best category combinations for these variables.

Variables with different ways of binning can lead to differences in selecting indicators because, within the same variable, different category combinations will result in different values of the c-statistic or adjusted R square. These are essential statistical criteria in the “MaxC” or “MaxR” forward stepwise approaches, which select indicators based on small improvements in these criteria. As a result, binning with different methods can lead to the selection of different indicators in targeting tools, which ultimately affects accuracy. This can be seen in comparing the WOE (Logit) and Scorecard (Logit) methods: despite both applying the same “MaxC” stepwise approach and using Logit models, their targeting tools differ in one indicator, and two to three indicators have different binning patterns. Consequently, their accuracy differs.

Additionally, for the Scorecard (Logit) and WOE (Logit) methods, the c-statistic value is also an important criterion for assessing the accuracy of targeting tools. We observe that variables with higher IVs often⁴ tend to achieve higher c-statistic values. Since the WOE (Logit) method maximizes the IV of indicators during the binning process, numeric variables with a wide range of values binned in this way often have higher c-statistic values compared to those binned using cross-tabulation. While this disparity is minor, it can accumulate when targeting tools contain many such indicators. As a result, this difference may contribute to the accuracy gap between Scorecards and WOE tools.

⁴ This trend weakens and not always correct when differences in IVs are very minimal.

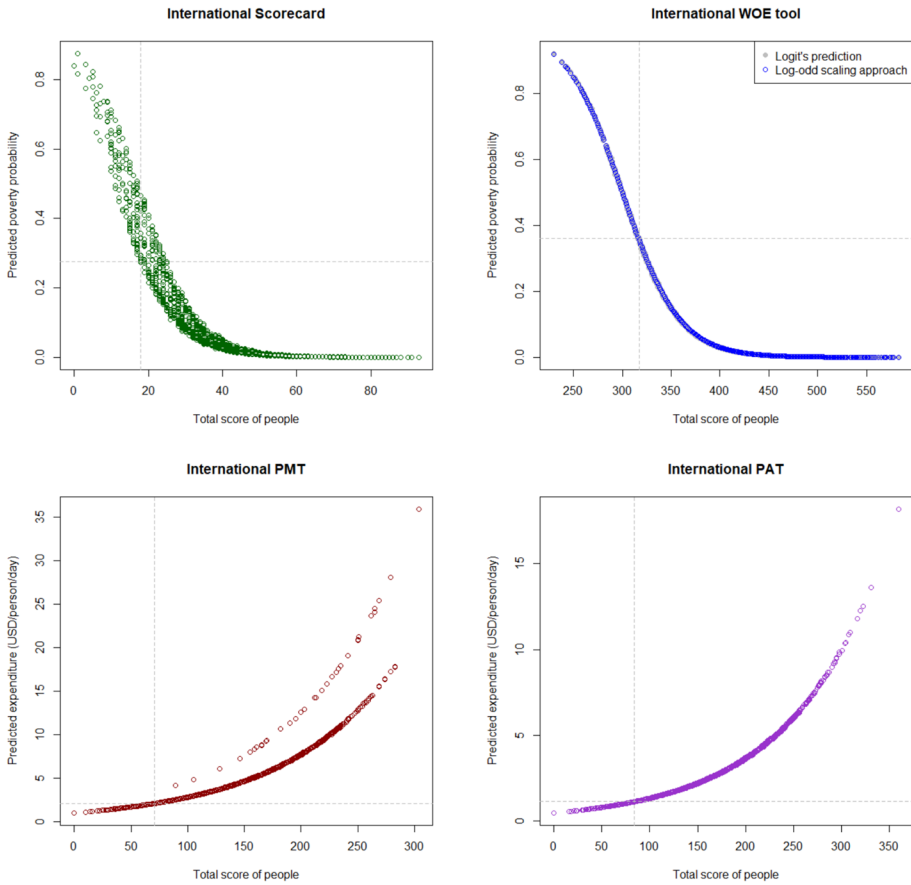


Fig. 1 The association between the total score and predicted outcomes at person level by international targeting tools

Another factor that may contribute to differences in indicator selection and targeting accuracy is the nature of the variables used in the regressions. The WOE (Logit) method employs numeric values transformed using the WOE technique, whereas other methods rely on categorized variables. However, we cannot definitively conclude that using WOE-transformed variables directly impacts accuracy or indicator selection, as the indicators used in the four targeting tools differ, potentially introducing bias.

An observation we can confirm is that regressions using categorized variables sometimes encounter issues with coefficient monotonicity. The Scorecard (Logit) and PMT (OLS) methods use categorical variables, so each category has its own coefficient. This can lead to situations where a categorical variable might help the model achieve the highest c-statistic or highest adjusted R-square but cannot be selected because their coefficients are non-monotonic.⁵ In such cases, we do not drop these variables; instead, we recategorize them to achieve monotonic coefficients and rerun the stepwise process. Whether the

⁵ A variable with non-monotonic coefficients leads to inconsistencies in the scoring of responses within a question.

re-selected indicator remains the same or not, this change often results in a model with a lower *c*-statistic or adjusted R-square. This issue arises when selecting indicators such as the number of dependents and number of agricultural laborers for the national Scorecard, and the highest education level for the international Scorecard. The PMT (Quantile) method faces similar challenges, so to maintain consistent coefficients across categories, we need to select quantiles that result in lower *c*-statistic values. In contrast, the WOE (Logit) method does not face this issue because WOE variables are numeric, and each variable has only one coefficient.

The third difference between methods is the regression algorithms used. However, we do not consider this factor as a main reason for the accuracy improvement seen in the WOE (Logit) method. Existing studies show that the Logit model sometimes outperforms OLS or Quantile regressions, but at other times, Logit models are less effective (Linh & Baulch, 2011; Schreiner, 2015; Zeller et al., 2006). This pattern is also observed in this study. For instance, in identifying poor people, the national Scorecard achieves higher accuracy than the national PMT and PAT, while the international Scorecard performs worse than the international PMT and PAT. In some cases, both the Logit and OLS regressions yield similar targeting accuracy (Etang et al., 2022). Additionally, although both the WOE (Logit) and Scorecard (Logit) methods use logistic regression, WOE tools still outperform Scorecards.

The final difference between the WOE (Logit) method and the other methods is the approach to transforming coefficients into scores. Figure 1 illustrates that the total scores converted by the WOE (Logit) method are more consistent with the predicted poverty probabilities than those converted by the Scorecard (Logit) method. This suggests that the Scorecard (Logit) method may inadvertently lower its accuracy, as many individuals with similar or even higher predicted poverty rates at the selected cut-off score may still be excluded. The magnitude of the total score could be a contributing factor to this issue. However, the score transformation approaches of the PMT (OLS) and PMT (Quantile) methods are similar, with large total score magnitudes. While the total scores in the PMT (Quantile) method align with individuals' expenditures, the total scores in the PMT (OLS) method do not. As a result, both the PMT (OLS) and PMT (Quantile) methods generally perform worse than the WOE (Logit) method. Nevertheless, concluding that the score transformation approach of the WOE (Logit) method is superior to others could be biased, as the targeting tools constructed by the four methods use different indicators, and the magnitude of the total scores varies across the methods.

This discussion highlights the differences between methods and the potential outcomes resulting from these differences. Based on these outcomes, we can only conclude that using both the WOE binning technique and WOE variables leads to differences in binned and selected variables compared to other methods. We cannot provide precise quantitative assessments of accuracy improvements for individual techniques within the WOE (Logit) method because the necessary conditions for a fair assessment are not met, and this is beyond the scope of the study. However, this presents an interesting research gap to explore in the future.

4 Conclusion

This study explores the feasibility and efficacy of using the WOE (Logit) method to develop targeting tools. To achieve this, we compare the accuracy of the WOE (Logit) method with that of other powerful and common methods, such as the Scorecard (Logit),

PMT (OLS), and PMT (Quantile) methods. For a fair comparison, we use the coverage rate to compare the accuracy between targeting tools while holding the predicted poverty rate.

The results indicate that targeting tools constructed using the WOE (Logit) method achieve higher accuracy than those developed by other methods. Depending on the national or international poverty line, the WOE (Logit) method can improve accuracy in identifying poor households by approximately 1.9–5.9 percentage points of coverage rate compared to other methods, while this accuracy improvement ranges from 1.5 to 3.3 percentage points in identifying the poor at the person level. This small improvement is meaningful on a national scale. For instance, according to the updated report by the Committee for Ethnic Minority Affairs of Vietnam (CEMA, 2020), Vietnam has 745,441 poor households among ethnic minorities. If the Vietnamese government aims to target all of these poor households, applying the WOE (Logit) method could potentially increase the number of correctly identified poor households by 14,163 to 43,981 compared to other methods.

This study introduces an additional approach for governments seeking to develop accurate targeting tools. While the WOE (Logit) method may seem complex in explanation, it is neither new nor difficult to implement, as it is commonly used in credit scoring. Various accessible resources, guidelines, and software packages are available for popular and powerful tools such as R, SAS, and Python. These packages allow users to construct targeting tools using the WOE (Logit) method in just a few simple steps.

The first limitation of this study is that it focuses solely on testing whether applying the WOE (Logit) method—a comprehensive approach for constructing credit risk scorecards, including techniques such as binning categories by WOE and IV, using WOE variables, applying logistic regression, and using the score-scaling approach of doubling the poverty odds—improves the accuracy of targeting tools. Therefore, the study can only conclude that the application of all these techniques together improves the accuracy of WOE tools. The individual contributions of each component technique to the accuracy improvement remain unclear, as many conditions are not controlled for a fair comparison, making any conclusion potentially biased. Nevertheless, this is a research gap that should be explored in the future if any researcher intends to apply any single technique in the WOE (Logit) method to the PMTs.

A second limitation is that the WOE (Logit) method has only been applied to ethnic minority communities in Vietnam. However, targeting tools developed using this method have been effective in identifying the poor at both the household and person levels among the 53 ethnic minority groups, which vary in characteristics, living regions, poverty rates, and levels of poverty (poor, moderately poor, or extremely poor) (CEMA, 2020). Therefore, the WOE (Logit) method should also be applicable for constructing targeting tools to identify the poor among the ethnic majority, which includes only the Kinh ethnic group, in Vietnam. We recommend that the comparative analysis of the WOE (Logit) method presented in this paper be repeated for other countries and data sets. If results are similar to those presented here, the WOE (Logit) method can be applied to poverty scoring in those countries.

Appendix

See Tables 8, 9 and 10.

Table 8 National Scorecard constructed by the Scorecard (Logit) method

Indicator	Value	Points	Score
1. How much did you pay for your electricity bill last month?	A. Less than 50,000 VND	0	
	B. 50,000 VND—99,000 VND	3	
	C. 100,000 VND—199,000 VND	7	
	D. 200,000 VND—349,000 VND	11	
	E. 350,000 VND or more	19	
2. How many members of the household work as non-agricultural laborers (includes side jobs)?	A. One or none	0	
	B. Two	7	
	C. Three	12	
	D. Four or more	17	
3. How many members does the household have?	A. Less than four	9	
	B. Four	5	
	Five	3	
	Six	2	
	Seven or more	0	
	A. Yes	0	
4. Was the household included in the anti-poor program in the previous year (2017)?	B. No	5	
	A. Less than two hectares	0	
5. What is the total land area owned by the household?	B. Two hectares or more	6	
	A. Yes	10	
6. Is anyone in the household working away from home?	B. No	0	
	A. None	12	
7. How many members of the household work in the agricultural sector as their main job?	B. One	9	
	A. Two, Three or Four	4	
	B. Five or more	0	

Table 8 (continued)

Indicator	Value	Points	Score
8. How many members of the households are dependents?	A. None	8	
	B. One or two	4	
	C. Three	1	
	D. Four or more	0	
9. Does the household own a cow, buffalo, or horse?	A. No	0	
	B. Yes	6	
10. How many motorbikes does the household own?	A. None	0	
	B. One	3	
	C. Two	6	
	D. Three or more	9	
		Total score	

Table 9 National WOE tool constructed by the WOE (Logit) method

Indicator	Value	Points	Score
1. How much did you pay for your electricity bill last month?	A. Less than 40,000 VND	17	
	B. 40,000 VND—89,000 VND	30	
	C. 90,000 VND—159,000 VND	43	
	D. 160,000 VND—239,000 VND	57	
	E. 240,000 VND or more	74	
2. How many members of the household work as non-agricultural laborers (includes side jobs)?	A. One or none	12	
	B. Two	40	
	C. Three	57	
	D. Four or more	65	
3. How many members does the household have?	A. Less than four	44	
	B. Four	37	
	C. Five	31	
	D. Six	30	
4. Was the household included in the anti-poor program in the previous year (2017)?	E. Seven or more	22	
	A. Yes	24	
	B. No	45	
	A. Less than 30	32	
	B. 30 to 39	35	
5. How many poultry does the household own?	C. 40 to 59	53	
	D. 60 or more	74	
	A. Yes	71	
6. Is anyone in the household working away from home?	B. No	32	
	A. None	26	
7. How many motorbikes does the household own?	B. One	32	
	C. Two	51	
	D. Three or more	64	

Table 9 (continued)

Indicator	Value	Points	Score
8. Does the household own a cow, buffalo, or horse?	A. No	33	
	B. Yes	58	
9. How many members of the household work in the agricultural sector as their main job?	A. None	68	
	B. One	50	
	C. Two or three	31	
	D. Four	24	
	E. Five or more	15	
10. How many members of the households are dependents?	A. None	53	
	B. One	40	
	C. Two	34	
	D. Three	20	
	E. Four or more	12	
	Total score		

Table 10 National PMT and PAT constructed by the PMT (OLS) and PMT (Quantile) methods, respectively

Indicator	Category	National PMT		National PAT	
		Points	Score	Points	Score
1. How much did you pay for your electricity bill last month?	A. Less than 50,000 VND	0		0	0
	B. 50,000 VND—99,000 VND	9		10	10
	C. 100,000 VND—199,000 VND	21		23	23
	D. 200,000 VND—349,000 VND	35		38	38
	E. 350,000 VND or more	50		51	51
2. How many members of the household work in the agricultural sector as their main job?	A. None	71		58	58
	B. One	49		39	39
	C. Two or three	25		18	18
	D. Four	6		3	3
	E. Five or more	0		0	0
3. How many members of the households are dependents?	A. None	58		53	53
	B. One	38		35	35
	C. Two	23		24	24
	D. Three	9		12	12
	E. Four or more	0		0	0
4. Was the household included in the anti-poor program in the previous year (2017)?	A. Yes	0		0	0
	B. No	23		17	17
5. How many members of the household work as non-agricultural laborers (includes side jobs)?	A. One or none	0		0	0
	B. Two	14		19	19
	C. Three	27		34	34
	D. Four or more	30		42	42
6. What is the highest level of education in the household?	A. Less than high school	0		0	0
	B. High school	3		0	0
	C. Above high school	28		29	29

Table 10 (continued)

Indicator	Category	National PMT		National PAT	
		Points	Score	Points	Score
7. Is anyone in the household working away from home?	A. Yes	24		21	
	B. No	0		0	
8. What is the total land area owned by the household?	A. Less than two hectares	19		14	
	B. Two hectares or more	0		0	
9. Does the household own a gas or electric stove?	A. Yes	17		11	
	B. No	0		0	
10. Does the household own a washing machine?	A. Yes	26		24	
	B. No	0		0	
	Total score			Total score	

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11205-025-03576-z>.

Acknowledgements The authors gratefully acknowledge data set supported from the General Statistics Office of Vietnam (GSO). The authors are solely responsible for any errors.

Authors contribution Be Thanh Duong: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization, Project administration. Orkhan Sariyev: Conceptualization, Methodology, Validation, Writing - Review & Editing, Supervision, Project administration. Manfred Zeller: Conceptualization, Methodology, Validation, Writing - Review & Editing, Supervision, Project administration.

Funding Open Access funding enabled and organized by Projekt DEAL. This work is supported by the German Academic Exchange Service (Deutscher Akademischer Austauschdienst, DAAD) as part of the IPPAE Ph.D. program at the University of Hohenheim.

Declarations

Conflict of interest No potential conflict of interest was reported by the authors.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Alsharkawi, A., Al-Fetyani, M., Dawas, M., Saadeh, H., & Alyaman, M. (2021). Poverty classification using machine learning: The case of Jordan. *Sustainability*, *13*, 1412. <https://doi.org/10.3390/SU13031412>
- Anderson, R. (2007). *The credit scoring toolkit: theory and practice for retail credit risk management and decision automation*. Oxford University Press. <https://archive.org/details/TheCreditScoringToolkitTheoryAndPracticeForRetailCreditRiskMgmt.DecisionAutomati>
- Attanasio, O., Sosa, L. C., Medina, C., Meghir, C., Posso-Suárez, C. M., & Taboada, B. (2023). The long-term effects of conditional cash transfer programmes: Evidence from Colombia. *Public Economics*. <https://voxdex.org/topic/public-economics/long-term-effects-conditional-cash-transfer-programmes-evidence-colombia>
- Baulch, B. (2002). *Poverty monitoring and targeting using ROC curves: examples from Vietnam* (161). Art. 161. <https://opendocs.ids.ac.uk/opendocs/bitstream/handle/20.500.12413/3940/Wp161.pdf>
- Burchi, F., Scarlato, M., & d'Agostino, G. (2018). Addressing food insecurity in Sub-Saharan Africa: The role of cash transfers. *Poverty and Public Policy*, *10*(4), 564–589. <https://doi.org/10.1002/POP4.233>
- CEMA, Unicef, & IRC. (2015). *Multidimensional child poverty of ethnic minority children: Situation, Dynamics and Challenges*. https://vietnam.un.org/sites/default/files/2019-08/Multidimensional%20child%20poverty%20in%20ethnic%20minority%20areas-%20egl%20UBDT_0.pdf
- CEMA. (2020). *Summary Results of Survey on the Socio-Economic Status of 53 ethnic minorities in 2019*. https://thongkedantoc.vn/static/media/files/file/2021_05_31/940_1622453704_43060b4adc806b5f.pdf
- IRIS Center. (2005). *Note on Assessment and Improvement of Tool Accuracy* (pp. 1–9). The IRIS Center. https://www.academia.edu/114542422/Note_on_Assessment_and_Improvement_of_Tool_Accuracy
- Chen, S., & Schreiner, M. (2009). *Simple Poverty Scorecard: Poverty-Assessment Tool Vietnam* (p. 159). https://www.simplepovertyscorecard.com/VNM_2006_ENG.pdf
- Chen, K., Zhu, K., Meng, Y., Yadav, A., & Khan, A. (2020). Mixed credit scoring model of logistic regression and evidence weight in the background of big data. *Advances in Intelligent Systems and Computing*, *940*, 435–443. https://doi.org/10.1007/978-3-030-16657-1_40
- de Maîté, K., & L. (2012). Conditional cash transfers in Latin America. *Reflets Et Perspectives De La Vie Economique, Tome L, I*(2), 5–18. <https://doi.org/10.3917/rpve.512.0005>

- Efron, B., & Tibshirani, R. J. (1994). *An introduction to the bootstrap*. CRC press. https://cindy.informatik.uni-bremen.de/cosy/teaching/CM_2011/Eval3/pe_efron_93.pdf
- Etang, A., Ervin, P., & Namara, S. (2022). A proxy means test for targeted social protection programs in Sudan. In *Policy Research Working Paper* (WPS 10241; Policy Research Working Paper, Vol. WPS10241). <http://documents.worldbank.org/curated/en/0993344411302240381>
- General Statistics Office. (2019). *Result of the Vietnam household living standards survey 2018* (T. L. Vu, A. K. Do, D. C. Nguyen, T. Q. Nguyen, & T. D. Lo, Eds.). Statistical Publishing House.
- Grosh, M. E., Baker, J. L., World, T., & Washington, B. (1995). Proxy means tests for targeting social programs. *Proxy Means Tests for Targeting Social Programs*. <https://doi.org/10.1596/0-8213-3313-5>
- Hand, D. J. (2006). Classifier technology and the illusion of progress. *Statistical Science*, 21(1), 1–15. <https://doi.org/10.1214/0883342306000000060>
- Kambuya, P. (2017). *Better Model Selection for Poverty Targeting through Machine Learning: a Case Study in Thailand*. http://ethesisarchive.library.tu.ac.th/thesis/2017/TU_2017_5904040077_8907_8836.pdf
- Kullback, S. (1997). *Information Theory and Statistics*. Dover Publications, Inc. <https://books.google.com/books?hl=vi&lr=&id=luHcCgAAQBAJ&oi=fnd&pg=PR13&dq=Information+Theory+and+Statistics,+kullback&ots=G2QAKnQ-PT&sig=59qZC074LDBcqMjGMvU4W408e6o>
- Leite, P. G., & George, P. (2014). Effective targeting for the poor and vulnerable. In *Safety Nets, Technical Note* (Vol. 6). World Bank Group. <https://documents.worldbank.org/en/publication/documents-reports/documentdetail/747591468125685125/effective-targeting-for-the-poor-and-vulnerable>
- Linh, V., & Baulch, B. (2011). Assessing alternative poverty proxy methods in rural Vietnam. In *Oxford Development Studies*, 39(3), 341–362. <https://doi.org/10.1080/13600818.2011.599207>
- Lund, B. (2016). *Weight of Evidence Coding and Binning of Predictors in Logistic Regression*. 16.
- McBride, L., & Nichols, A. (2015). *Improved poverty targeting through machine learning: An application to the USAID Poverty Assessment Tools*. https://www.econthatmatters.com/wp-content/uploads/2015/01/improvedtargeting_21jan2015.pdf
- Nguyen, V. C., & Lo, D. (2016). *Testing Proxy Means Tests in the Field: Evidence from Vietnam* (p. 35). Munich Personal RePEc Archive. <https://mpira.ub.uni-muenchen.de/id/eprint/80002>
- Nguyen, C. V., & Tran, D. T. (2018). Proxy means tests to identify the income poor: Application for the case of Vietnam. In *Journal of Asian and African Studies*, 53(4), 571–592. <https://doi.org/10.1177/0021909617709486>
- Ohlenburg, T., Pinxten, J., Fricke, D., & Caccioli, F. (2022). *Fewer Questions, More Answers Truncated Early Stopping for Proxy Means Testing* (WPS 10255; Policy Research Working Paper). <http://documents.worldbank.org/curated/en/099339212072228483>
- Pomazanov, M. (2023). Continuous transformation of nonmonotonic factors of the scoring model increasing its discriminating power. *Procedia Computer Science*, 221, 237–244. <https://doi.org/10.1016/J.PROCS.2023.07.033>
- Ravallion, M. (2007). How Relevant is Targeting to the Success of an Antipoverty Program? *Policy Research Working Paper*, 4385, 34. <https://documents1.worldbank.org/curated/en/444201468137704822/pdf/wps4385.pdf>
- Rawlings, L. B., & Rubio, G. M. (2003). *Evaluating the Impact of Conditional Cash Transfer Programs Lessons from Latin America* (3119; Policy Research Working Paper). <http://econ.worldbank.org>.
- Schreiner, M. (2006). *Simple Poverty Scorecard: Poverty-Assessment Tool Bangladesh* (p. 47). https://www.simplepovertyscorecard.com/BGD_2000_ENG.pdf
- Schreiner, M. (2012). *Simple Poverty Scorecard: Poverty Assessment Tool Afghanistan* (p. 117). https://www.simplepovertyscorecard.com/AFG_2007_ENG.pdf
- Schreiner, M. (2015). A comparison of two simple, low-cost ways for local, pro-poor organizations to measure the poverty of their participants. *Source: Social Indicators Research*, 124(2), 537–569.
- Schreiner, M. (2022). *How to Transform Regression Coefficients to Integer Weights So Scores Go from 0 to 100*. 5. <https://microfinance.com/English/Papers/How-to-Present-Logit-Weights-So-Non-Specialists-Understand.pdf>
- Schreiner, M. (2023). *Simple Poverty Scorecard Tool Cambodia*. 1–127. <https://www.simplepovertyscorecard.com/KHM-2019-ENG.pdf>
- Schreiner, M. (2010). A simple poverty scorecard for Pakistan. *Article Journal of Asian and African Studies*, 45(3), 326–349. <https://doi.org/10.1177/0021909609353579>
- Schreiner, M., Matul, M., Pawlak, E., & Kline, S. (2014). *Poverty Scorecards: Lessons from a Micro-lender in Bosnia-Herzegovina*. <https://doi.org/10.1002/pop4.86>
- Sharma, D. (2011). Evidence in favor of weight of evidence and binning transformations for predictive modeling. *SSRN Electronic Journal*. <https://doi.org/10.2139/SSRN.1925510>

- Siddiqi, N. (2006). *Credit Risk Scorecards: Developing and Implementing Intelligent Credit Scoring*. John Wiley & Sons, Inc. <https://www.wiley.com/en-au/Credit+Risk+Scorecards%3A+Developing+and+Implementing+Intelligent+Credit+Scoring-p-9781119201731>
- Siddiqi, N. (2016). *Intelligent Credit Scoring: Building and Implementing Better Credit Risk Scorecards* (Second Edition). John Wiley & Sons, Inc. <https://onlinelibrary.wiley.com/doi/book/https://doi.org/10.1002/9781119282396>
- Skoufias, E., Diamond, A., Vinha, K., Gill, M., & Dellepiane, M. R. (2020). Estimating poverty rates in subnational populations of interest: An assessment of the Simple Poverty Scorecard. *World Development*, 129, 104887. <https://doi.org/10.1016/J.WORLDDEV.2020.104887>
- USAID. (2006). *USAID/IRIS Tool for Vietnam* (p. 5). United State Agency International Development. https://www.academia.edu/114543070/USAID_IRIS_Tool_for_Vietnam
- Wodon, Q. T. (1997). Targeting the poor using ROC curves. *World Development*, 25(12), 2083–2092. [https://doi.org/10.1016/S0305-750X\(97\)00108-3](https://doi.org/10.1016/S0305-750X(97)00108-3)
- Zeller, M., Gabriela Alcaraz, V., & Johannsen, J. (2005). *Developing and testing poverty assessment tools: Results from accuracy tests in Bangladesh*. https://www.researchgate.net/publication/320288550_Developing_and_testing_poverty_assessment_tools_Results_from_accuracy_tests_in_Bangladesh
- Zeller, M., Houssou, N., Alcaraz, G. V., Schwarze, S., & Johannsen, J. (2006). Developing poverty assessment tools based on principal component analysis: Results from Bangladesh, Kazakhstan, Uganda, and Peru. In *International Association of Agricultural Economists Conference, Gold Coast, Australia, August 12–18* (Vol. 4592175, p. 24).

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.