How accurate is Participatory Wealth Ranking (PWR) in targeting the poor?

A case study from Bangladesh

Manfred Zeller*, Joseph Feulefack, and Andreas Neef***

*Institute for Agricultural Economics and Social Sciences in the Tropics and Subtropics,

University of Hohenheim, Stuttgart, Germany

** International Institute for Tropical Agriculture (IITA), Cameroon

Contributed paper selected for the International Association of Agricultural Economists

Conference, Gold Coast, Australia,

August 12-18, 2006

Corresponding author:

Prof. Dr. Manfred Zeller
Institute of Agricultural and Socioeconomics in the Tropics and Subtropics, University of Hohenheim
Schloss, Osthof (490a)
70599 Stuttgart
GERMANY
Tel. +49-711-4592175
Fax: +49-711-4592794
Email: manfred.zeller@uni-hohenheim.de

Abstract

PWR is a participatory poverty assessment method that uses the ratings of local reference groups concerning the relative poverty status of households in their community. This paper assesses the accuracy of PWR in predicting absolute (income) poverty, and compares PWR with three other poverty assessment methods. Using a village census in 8 villages located in three of the six divisions of Bangladesh, 1660 households have been scored using the PWR method. A randomly selected subsample of 320 households was interviewed with a questionnaire employing the Living Standard Measurement Survey (LSMS) method. The data allow the identification of households that have per-capita expenditures below the international poverty line of 1 dollar a day. Our results show that calibrated PWR scores can achieve an accuracy of 70 to 79 percent, i.e. up to 8 out of 10 households are correctly predicted as to whether they live in extreme poverty or not. As expected, the so-called Total Accuracy of PWR is higher if its scores are calibrated at lower geographical level, and highest if calibrated at the community level. For the case of Bangladesh, the results confirm the accuracy of PWR as a poverty targeting method for development policies and projects if used at the community level.

JEL subject codes:

I3 Welfare and Poverty, C8 Data Collection and Data Estimation Methodology

Keywords:

Participatory wealth ranking, poverty targeting, Bangladesh, LSM

1 Background and aim of the study

At the beginning of this Millennium, the UN adopted the Millennium Development Goals (MDG). The first of these goals seeks to halve the proportion of people who are living in extreme poverty (i.e. below 1 dollar a day per capita at purchasing power rates) by 2015 (UNDP, 2003). Another initiative was taken at the Microcredit Summit in 1997, Washington, where participants from many countries committed themselves to "join forces to ensure that 100 million poor families receive quality microfinance services by the year 2005" (UNESCO, 1997). In 2003, the US Congress enacted a law with the commitment to have at least half of Microenterprise funds reach those living in conditions of extreme poverty. Since then, USAID has been working with experts to develop accurate and cost-effective poverty measurement tools. In this respect, the Microcredit Summit Campaign Executive Committee approved the creation of a Microcredit Summit Poverty Measurement Tool Kit (PMTK). PWR and the Housing Index are the two first tools to be considered for the certification process (Microcredit Summit Campaign 2005).

Notwithstanding, the question as to which extent PWR can confidently be used to target the extreme poor and to assess ex-post the poverty level of clients of development programs or projects operating remains an open research issue. PWR measures relative poverty, not absolute poverty defined by the international standard of 1 dollar a day. The ability of PWR, for instance, to accurately inform about absolute poverty, its practicality of implementation and even its cost-effectiveness are still not sufficiently explored by socioeconomic research. Therefore, institutions using PWR or other tools, such as housing index or the CGAP Poverty Assessment Tool (Henry et al., 2003), lack sound empirical evidence on whether their tool is able to accurately assess whether new applicants into their poverty-targeted program live below or above the 1 dollar-a-day poverty line or not. This lack of information, however, is a major hindrance in improving targeting efficiency so as to make progress towards the MDG goals or those set by the Microcredit Summit. This paper seeks to assess the accuracy of PWR in predicting absolute poverty, and compares it with the accuracy of three other tools: (1) the Microfinance Poverty Assessment Tool (PAT) of the Consultative Group to Assist the Poorest (CGAP), (2) the subjective assessment of the household's poverty status by the interviewer also called 'visual impression'(VI), and (3) the Ladder of Life (LL). We examine three questions[i]:

(1) Is there any significant correlation between the poverty measures and per-capita daily expenditures, measured with the method used in the Living Standard Measurement Survey (LSMS) developed by the World Bank?
(2) How accurate is PWR (and the other three tools) in predicting a household being below or above the international poverty line of one-dollar a day?
(3) How does the accuracy performance of PWR change if the scores are calibrated at higher different geographical levels (e.g. district instead of community)?

---

[i] The data were collected by the survey firm DATA in Bangladesh in the scope of the project Developing Poverty Assessment Tools, which is carried out by the IRIS Center, University of Maryland, and funded by the United States Agency for International Development (USAID) under the Accelerated Microenterprise Advancement Project (AMAP) (Contract No. GEG-I-02-02-00029-009). We gratefully acknowledge the source of the data. The cleaning and aggregation of the data (including the daily per-capita expenditures) were carried out at the Institute of Rural Development, University of Goettingen, Germany. We are thankful to Gabriela Alcaraz and Stefan Schwarze for their helpful suggestions. We are also grateful for comments received from Thierry van Bastelaer, Christian Grootaert, Kate Druschel, and Laura Foose on a previous version of our analysis regarding PWR contained in Zeller et al. (2005).

2 Methods

2.1 Design of field research

The PWR was carried out in 8 villages from 4 districts of Bangladesh. The survey districts are: Barisal in the south at the Bay of Bengal, Dhaka in the centre, and Rajshahi in the northwest. The field research comprised an LSMS-type household expenditure survey (Grosh and Glewwe, 1998) and a PWR. The PWR covered all 1660 households (census method) in the eight selected communities. For the expenditure survey, 40 households in each village (i.e. n=320) were randomly selected.

Participatory Wealth Ranking is a method whereby communities define themselves who the poorest or the better-off are. Quoting Gibbons et al., 1999, p.43: “We are interested in peoples’ own ideas about poverty. We want them to tell us what they think and to tell us who in their village are very poor, poor or better off”. The PWR begins with a community-wide meeting convened by the facilitation team. After discussing the meaning and understanding of poverty in the local context, the people draw a map of all the households in the village and fill a card with the name of each household. Three reference groups are then formed in each ranking section, i.e. the hamlet. In Bangladesh, only women were asked to join the groups. After filling the cards, each reference group then meets separately and sorts the household cards into piles according to the living standard on a continuum from high to low. Next comes the crosschecking whereupon the results of the ranking done by the three reference groups are brought together and the piles are scored. Scores are calculated according to the number of piles used by participants, using the following formula: *Score of reference group =[100/(number of piles)] x pile number.*

For instance, if there are four piles, then the poorest pile (number 4) will score 100 by using the formula (100/4 x 4 = 100), and the richest pile (number 1) will score 25 by using the formula (100/4 x 1 = 25). The final score of each household is the average of the scores given by the three reference groups. Thus, the PWR methodology used in the field research closely followed the one developed by Gibbons et al. (1999). To ensure a consistent implementation of the PWR process, the facilitators were trained in a PWR course at the Bangladesh Academy for Rural Development in Comilla, organized by the Microcredit Summit in February, 2004.

2.2 Verification of data consistency

Data were entered and cleaned with SPSS (Statistical Package for Social Sciences). Following Gibbons et al. (1999), three main cases can be distinguished when evaluating the internal consistency of the PWR scores. The first case of highly consistent scores is given if the scores by the three reference groups do not deviate more than 25 score points. The second case is defined by a deviation of above 25, but below 50 points, whereas the third case of inconsistent scores shows deviations of 50 points or more. Following the procedure for consistency checks proposed by Gibbons et al. (1999), we remove the 27 households in two hamlets in Chak Shadu where we observe a high number of inconsistent scores. Thus, the sample size for accuracy analysis drops from 320 to 293 households. We conclude that the overall quality of the remaining data is within the acceptable range as defined by the PWR manual by Gibbons et al. (1999).

2.3 Accuracy analysis

Accuracy is the degree of conformity with a benchmark that is considered to be the "truth". The benchmark used in our case is the LSMS-type daily per capita expenditure measure, coupled with the absolute poverty line of 1 dollar a day measured at the purchasing power parity rate. At time of survey in March 2004, 1 US-Dollar was equivalent in purchasing power to 23.18 Taka, the currency in Bangladesh (Zeller et al., 2004). Households with per-capita expenditures below this international poverty line are rated as very-poor (VP), otherwise not very-poor (NVP). Using this poverty line, we find 96 households (or 32.8%) of the sample of 293 households to be very-poor. The research task consists of determining the accuracy of a tool, for instance PWR, to correctly predict whether the household is very-poor (VP) or not very-poor (NVP).

2.4 Calculation of accuracy performance

PWR and the other three tools generate scores that rank households with respect to relative poverty. the score of PWR ranges between 0 and 100. Here, higher PWR scores indicate a higher degree of poverty. The CGAP PAT, using principal component analysis, computes a N(0,1) distributed poverty index as an aggregate score from a set of indicators (Zeller et al., 2006). Higher values of the poverty index indicate lower relative poverty. The PAT is a recently invented approach of measuring relative poverty with the help of a composite index derived by Principal Component Analysis (PCA). We carry out the PCA analysis following the guidelines of the CGAP Microfinance Poverty Assessment Tool, which was primarily designed to measure the levels of well-being of clients of micro-finance institutions (Henry et al., 2003). The PAT can be used for the assessment of poverty outreach of all types of development institutions or projects that target the poor, such as in the areas of agriculture or social policy.

The scores for the Ladder of Life (LL) and for the Visual Impression by the interviewer ranges between 1 and 10, and 1 and 5 respectively. Each of the methods is briefly explained next. In the visual impression (VI) method, the interviewer subjectively rates the living standard of the respondents' household when applying a formal questionnaire concerning poverty indicators. After each section of the questionnaire, each covering a different dimension of poverty, the interviewer rates the poverty level of a household on a Likert scale from 1 to 5. The interview is conducted at the residence of the household so that the interviewer uses the information gained from observation as well. The Ladder-of-Life (LL) method uses a picture of a ladder with 10 steps. The respondent is informed that the lowest step represent the poorest in society whereas the highest step represents the richest. This picture is presented to the respondent who is invited to tell what step might best describe his or her own living standard. The following question is asked, “On which step of the ladder are you located today?” Hence, a higher step indicates less poverty.

In order to test the accuracy of a tool that measures relative poverty, a cut-off score is sought that maximizes the tools’ accuracy. For example, we may calibrate PWR at a cut-off score of 85. We then can calculate the accuracy of the decision rule that states that all households with scores of 85 or above are predicted as very-poor, and all households below that cut-off score are not very-poor. The poverty status predicted by the simulated decision rule of a tool is then compared with the “true” poverty status as determined by the poverty benchmark. Through a systematic simulation of alternative scores, a so-called BEST score can be found that maximizes the tools’ accuracy with respect to a predetermined criterion. We term this simulation method the BEST score method. It is consistently applied for all four tools.

IRIS (2005) distinguishes seven accuracy criteria. Total Accuracy is the percentage of households whose poverty status is correctly predicted by the tool. Poverty Accuracy, i.e. the accuracy among the very-poor, refers to the households correctly predicted as very-poor, expressed as a percentage of the total number of very-poor. Vice versa, Non-Poverty Accuracy refers to households correctly predicted as not very-poor, expressed as percentage of the total number of not very-poor. Undercoverage represents the error of predicting very-poor households as being not very-poor, expressed as a percentage of the total number of very-poor. Leakage reflects the error of predicting not very-poor households as very-poor, expressed as a percentage of the total number of very-poor. The Poverty Incidence Error (PIE) is defined as the difference between the predicted and the actual (observed) poverty incidence, measured in percentage points. Finally, the Balanced Poverty Accuracy Criterion (BPAC), is defined as: Poverty Accuracy minus the absolute difference between Undercoverage and Leakage, each expressed as a percentage of the total number of very-poor. When Undercoverage and Leakage are equal, the BPAC is equal to the Poverty Accuracy. BPAC is measured in percentage points (IRIS, 2005). A perfectly accurate tool has a PIE of 0% and a BPAC of 100%. For reasons of brevity, we restrict the presentation of results to only three criteria, namely the Balanced Poverty Accuracy Criterion (BPAC), the poverty incidence error (PIE), and the Total accuracy (TA). Moreover, we choose BPAC as the criteria for calibration of each of the four tools.

3 Results and discussion

3.1 Correlation of PWR and other tools with poverty benchmark

The results displayed in Table 1 show that PWR correlates less with per capita daily expenditures than it does with the VI. PWR’s correlation with the benchmark is also higher than that of the LL. This underlines the better harmony the benchmark shows with a local consensus form of ranking like PWR than with the subjective self-assessment such as the LL. This supremacy of reference group-based ranking over individual self-assessment was also reflected in the findings by Nga Nguyet Nguyen and Rama (cited from World Bank 2003, p. 120). They found that PWR’s correlation coefficient with the benchmark of 0.462 was higher than the one of 0.378 for the self-assessment. However, while PWR’s correlation coefficient was found to be significant, that of the self-assessment was not. This result suggests the extreme uncertainty involved in such subjective methods that rely on the self-assessment view of the respondent.

We find that the correlation between the benchmark and VI is higher than of the PWR. The interviewer is influenced by the responses to the questionnaire, which may bias his rating in favour of quantitative poverty indicators such as expenditures and value of assets that are asked during the interview. The external methods of assessment (PAT, VI) correlate better with the benchmark whereas the internal methods of assessment (PWR, LL) have correlation coefficients below 0.50. An insightful study by Häuser (2005) in Vietnam found that expenditure ranking correlated better with the PAT than with the self-assessment of the respondent, the latter being a method more comparable to the LL in our case. Similarly, in our study, we take note of the clustering of the external methods of poverty assessment on the one hand and that of internal poverty assessments on the other hand. Hence, internal methods appear to be more based on the personal judgement of local people regarding qualitative, more subjective indicators while the external methods may be more influenced by (monetary) indicators also found in the benchmark.

3.2 Comparison of accuracy performance measures of PWR and other tools

Figure 1 compares accuracy performance of the tools across three geographical levels. The PAT's curve is above all the other curves with a BPAC above 62%. Thus, irrespective of the geographical level, the PAT achieves the best accuracy performance. What are the levels of the BEST score for each of the tools? For the 'nation' level (i.e. represented by the full sample of 293 households), the PAT was has been calibrated with a BEST score of –0.589 so as to maximize BPAC. Hence, the PAT's decision rule to rate a household as VP or not is as follows: Households with a poverty index score of –0.589 or less are predicted as very-poor. Likewise, and again for the nation level, the BEST score for PWR is 86,67; for LL it is 2; and for VI, it is 2.11 on average for all the assessment. Though all the other tools have BPAC values around 50% at the 'nation' level, they are, however, clearly demarcated from each other at the community level such that the VI is always on top, i.e. the closest to the PAT's curve. PWR is second to the last with the lowest BPAC around 40% at the community level wherein the LL performs worst by scoring BPAC values below 30%.

A perfect tool has a PIE of 0 percent, indicating that the predicted poverty rate is the same than the observed one in the sample. In Figure 2, the PAT's curve is the closest to the zero level, except for the district where PWR does slightly better. PWR's curve is more regular than that of the remaining tools, i.e. indicating a higher robustness across geographic levels. Figure 3 presents the Total Accuracy for each of the four tools across the geographical levels. PAT and VI as external assessments outperform the self- assessment (LL) or group-based poverty assessment (PWR) tools at all levels.

3.3 Discussion

Whatever the geographical level, PAT predicted absolute poverty status better than the other low-cost tools when using BPAC as the accuracy criterion. PAT was followed by VI in the second position whereas PWR only came third. The ladder of life lagged behind. This result suggests the superiority of the PAT.

In terms of total accuracy, PWR has been outperformed by all the three other tools at any level. Nevertheless, this result may be specific to Bangladesh and cannot be generalised to other countries. Gibbons et al. (1999, p.39) compared PWR with the Housing Index and came to the stipulation that, "externally judged criteria, produced less accurate results in [their] working area, when compared to a local judgement of poverty." However, their research lacked per-capita daily expenditures as a 'true' benchmark reference for absolute poverty.

The results with respect to the Poverty Incidence Error (PIE) show that PAT most accurately predicts the observed poverty rate across the three geographic levels. Among the subjective tools, PWR is the least prone to misspecifications at the district and the nation level. At all geographic levels and for all four tools, inaccuracy in predictions among the very-poor was higher than those for the not very-poor.

Each of the four tools examined in this paper was found to significantly correlate with the poverty benchmark. PAT had the highest correlation coefficient of nearly 0.6. External methods of assessment correlated better with the poverty benchmark than internal methods involving judgements based on facts and experiences of the local people. Correlation coefficients were in general relatively far from one in absolute term, which may be due to the fact that the different tools do rank based on different dimensions of poverty. The patterns displayed by the BPAC curves in Figure 1 corroborate the correlation results. Actually, Figure 1 shows that in villages and districts, the higher the correlation with the benchmark, the upper the curve and, hence, the better the BPAC.

4 Conclusion

The analysis of accuracy of PWR and three other 'low-cost' poverty assessment tools have disclosed three main results. First, PWR's Total accuracy in the assessment of the aggregate poverty status at the BEST score tends most often to decrease as the area enlarges. This result is expected as PWR relies on group-based judgements within and concerning a specific community. Second, PAT has outperformed the other low-cost tools in terms of the three main accuracy and error measures: TA, PIE, and BPAC. Third, external tools (PAT and VI) showed higher correlation with the benchmark than the internal tools (PWR, LL) such that the difference in correlation coefficients among tools of the same type was relatively low. Given that PWR can correctly identify the majority of households i.e. nearly 7 out of 10 at the nation level and nearly 6 out of 10 among the very-poor; given that the wrongly predicted households could be due to the focus of the predicting standard (here LSMS-type per capita daily expenditures ranking) on monetary dimension of poverty while the PWR tool includes other aspects of the human behaviour (happiness, always welcoming, all the daughter are married, having friends to rely on, etc.) from difficult-to-measure dimensions of poverty; and given that PWR tool as a "consensus" form of ranking has relatively performed better in terms of the predicted poverty rate compared to the other tools examined (except PAT); these results suggests that one can rely on PWR as a complementary method to more established absolute poverty measures for targeting the poor at the community level. The process of PWR may also contribute to fulfil what is sought by IFAD (2001, p.20-21), namely that, "understanding the 'psycho-emotional environment' of the rural poor and their personal and familial perception and aspirations will contribute to the success or failure of rural development programmes”.

The PAT outperformed the other three tools at the national level for all the accuracy measures (BPAC, PIE, and TA) tested in this analysis. With respect to BPAC, the PAT is matchless: it outperforms any of the other three tools at each of the geographical levels. Given its superior accuracy performance, the PAT appears therefore as the most suitable tool (among the four tools tested in this paper) to assess ex-post the poverty outreach among clients of development policies and projects that aim to reach the poor with their services. The results further suggest that the PAT can also be used for targeting by development institutions such as microfinance institutions. An added advantage of the PAT – compared to the other three tools analyzed in this paper - is that the indicators and their weights can be documented such that an evaluator can compute a poverty index score after having asked information concerning the indicator from the program’s applicant. We recommend that further research on systematic comparison of different poverty assessment tools – compared with accepted benchmarks for absolute poverty- is undertaken so as to confirm or contradict our results. Such research will eventually improve the tools used by development practitioners and hence improve the targeting efficiency of development policy.

Table 1: Correlation of scores of poverty assessment tools with the poverty benchmark (n=293)

	Ladder of life	Poverty Assessment Tool	Participatory Wealth Ranking	Visual impression, community as reference	Per capita daily expenditures (Poverty benchmark)
Ladder of life today?	1				
Poverty Assessment Tool	0.632*	1			
Participatory Wealth Ranking	-0.586*	-0.700*	1		
Visual impression, community as reference	0.768*	0.730*	-0.721*	1	
Per capita daily expenditures	0.416*	0.588*	-0.425*	0.522*	1

Note: An asterisk denotes that the Pearson correlation coefficient is significant at the 1% level (2-tailed).

Figure 1: Balanced Poverty Accuracy Criterion (BPAC), by geographical level and tool

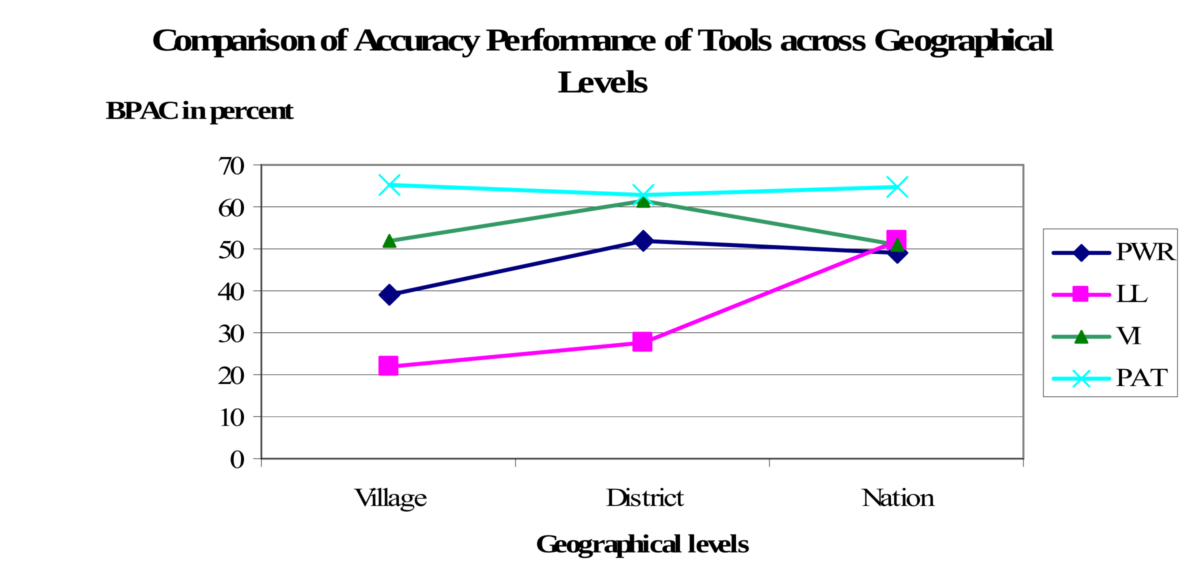
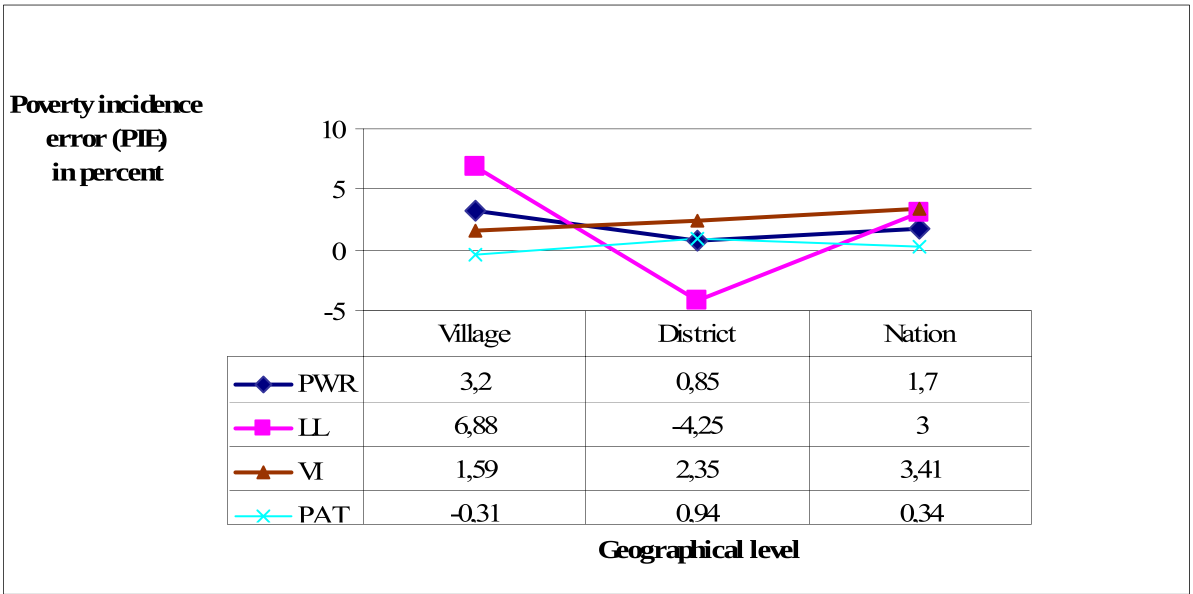
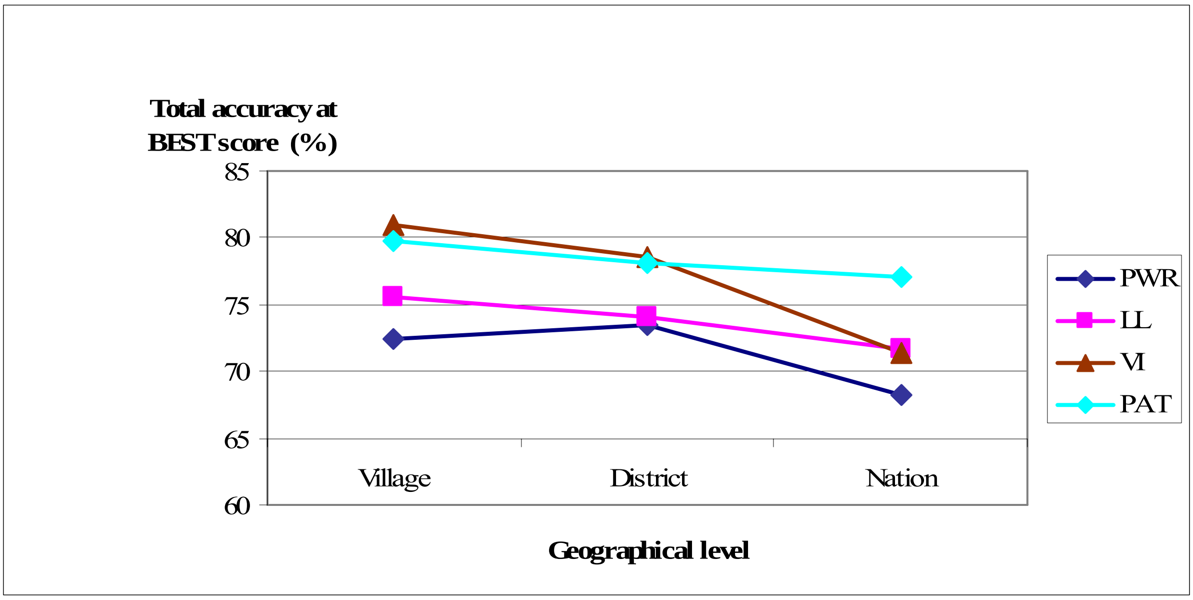


Figure 2: Poverty Incidence Error (PIE), by geographical level and tool



	Village	District	Nation
PWR	3,2	0,85	1,7
LL	6,88	-4,25	3
VI	1,59	2,35	3,41
PAT	-0.31	0.94	0.34

Figure 3: Total Accuracy (TA), by geographical level and tool



References

Gibbons D.S., Simanowitz A., Nkuna B., 1999. CASHPOR-SEF Operational

Manual, Cost-effective Targeting: Two Tools to Identify the Poor. Edited by Helen Todd. Produced by CASHPOR Technical Services. Sermban, Negri Sembilan, Malaysia.

Grosh, M, and Glewwe, P. 1998. “The World Bank’s Living Standard Measurement Study Household Surveys.” Journal of Economic Perspectives 12, 187-196.

Häuser, I., 2005. Measuring Poverty in Vietnam: Strengths and Weaknesses of Different

Indicators. Master Thesis. Institute of Agricultural Economics and Social Sciences in the Tropics and Subtropics, University of Hohenheim. Stuttgart-Hohenheim.

Henry, C., Sharma, M., Lapenu, C., Zeller, M. 2005. Microfinance poverty assessment tool. Technical Tools Series No. 5, September 2003. Published by the Consultative Group to Assist the Poor (CGAP) and The World Bank, Washington, D.C.

IFAD. 2001. Assessment of Rural Poverty, Latin America and the Caribbean. Edited by Benjamin Quijandria, Anibal Monares, Raquel Ugarte de Peňa Montenegro. Rome, Italy, 147 pp.

IRIS. 2005. Note on Assessment and Improvement of Tool Accuracy. Developing Poverty Assessment Tools Project. Mimeograph, Revised version from June 2, 2005. IRIS Center, University of Maryland, 9pp.

Microcredit Summit Campaign. 2005. Microcredit Summit Campaign Establishes Poverty Measurement Tool Kit. Newsletter of Microcredit Summit Campaign. Available at http://www.microcreditsummit.org/newsletter .

UNDP. 2003. Human Development Report 2003: Millennium Development Goals. Edited by Fukuda-Parr S., New York Oxford University Press.

UNESCO. 1997. Microcredit Summit. Position paper prepared by UNESCO (CAB-97/WS/2). 2-4 February 1997. Washington, DC.

World Bank. 2003. Poverty. Vietnam Development Report, Joint Donor Report to the

Vietnam Consultative Group Meeting, Hanoi, December 2-3. 147pp.

Zeller, M., Alcaraz, G.V., Johannsen, J. 2005. Developing and testing poverty assessment tools: Results from accuracy tests in Bangladesh. IRIS Center, University of Maryland, College Park, 100 pp (Available at http://www.povertytools.org).

Zeller, M., Sharma, M., Henry, C.J., Lapenu, C. 2006. An operational method to assess the poverty outreach performance of development policies and projects: Results of Case Studies in Africa, Asia and Latin America. World Development, 34 (3), 446-464.