A validity assessment of the Progress out of Poverty Index (PPI)™ for Rwanda

Sam Desiere[1*], Wytse Vellema[1], Marijke D'Haese[1]

[1] Department of Agricultural Economics, Ghent University, Coupure Links 653, B-9000 Ghent, Belgium
*Corresponding author: sam.desiere@ugent.be

Paper prepared for presentation at the EAAE 2014 Congress
'Agri-Food and Rural Innovations for Healthier Societies'

August 26 to 29, 2014
Ljubljana, Slovenia

Abstract
Development organisations need easy-to-use and quick-to-implement indicators to quantify poverty when requested to measure program impact. In this paper we assess the validity of the Progress out of Poverty Index (PPI)^{TM}, a country-specific indicator based on ten closed questions on directly observable household characteristics, by its compliance to the SMART criteria. Each response receives a pre-determined score, such that the sum of these scores can be converted into the likelihood the household is living below the poverty line. We focus on the PPI scorecard for Rwanda, which was validated using two national household surveys conducted in 2005/06 and 2010/11. The PPI is Specific, Measurable, Available cost effectively, and Timely available. Yet, its Relevance depends on the way it is used. Although it accurately distinguishes poor from non-poor households, making it a useful reporting tool, its limited sensitivity to changes in poverty status restricts its usefulness for evaluating the impact of development projects.
Key words: Progress out of Poverty Index (PPI)^{TM} , expenditure poverty, poverty measurement, impact evaluation, Rwanda

1 Introduction

Development programs with the objective of poverty alleviation want to target the poorest households, but lack resources, time and expertise to develop their own detailed poverty measures or conduct full-scale household surveys. Consequently, such development programs rely on standardized indicators to measure poverty and evaluate the impact of their program. Ideally, such indicators are designed according to the SMART criteria[1]: *Specific*, *Measurable*, *Available cost-effectively*, *Relevant* and *Timely available* (European Evaluation Network for Rural Development, 2014; Poister, 2008). The Progress out of Poverty Index (PPI)^{TM}, introduced by the Grameen Foundation, is promoted as a tool that can quantify the share of program participants living below the poverty line, assess the performance of the intervention among the poor and poorest, and track poverty levels over time.[2] By design, the PPI meets four of the five SMART criteria. It is Specific, Measurable, Available cost-effectively, and Timely Available. The relevance criterion, however, requires validation. Assessing this validity is the objective of this paper.

Background

In order to be able to more accurately value the merits and shortcomings of the PPI, we provide a brief overview of alternative ways to measure poverty, paying particular attention to the extent to which these measures comply with the SMART criteria.

Consensus appears to have been reached on the vision that poverty is multidimensional. However, such consensus does not exist on the best way to measure poverty, evidenced by the large and growing number of poverty indicators. The most frequently used poverty indices are income- or expenditure-based (M. Ravallion, 1996). Of these, perhaps the most well-known are the dollar-a-day extreme poverty line and the more generous two-dollar-a-day poverty line developed by Martin Ravallion, Datt, and Walle (1991) for the 1990 World Development Report. Households are considered poor when their income or total expenditure falls below a certain threshold. A downside of income and expenditure-based poverty indices is that data collection is

[1] Several definitions of SMART have been developed. We follow the definition as proposed by the European Evaluation Network for Rural Development of the European Commission (http://enrd.ec.europa.eu/en, accessed April 2014)
[2] www.progresssoutofpoverty.org

costly, extremely time-consuming and prone to measurement error (Beegle, De Weerdt, Friedman, & Gibson, 2012; Deaton, 1997). For example, the food expenditure part of the 2005/06 Rwanda Household Living Standard Survey counted 75 pages and required enumerators to visit each household 11 times (Schreiner, 2010b). Hence, expenditure-based poverty indicators are neither *Available cost effectively* nor *Timely available*, which are key SMART principles for successful implementation by development programs.

Additional shortcomings of income and expenditure based poverty measures are the difficulty of accurately defining a line below which people are poor and above which they are not (i.e. quantifying the poverty lines) and their inherently static nature (Carter & Barrett, 2006). Most income-based measures are static in nature as they measure only if a household (or individual) fails to meet minimum income levels to cross a predefined poverty line. Yet, some households – the so-called transient poor - live from an income below the poverty line at a particular moment in time, but have sufficient productive assets to escape poverty, while structurally poor households lack the resources to move out of poverty over time (Barrett et al., 2006a; Baulch & Hoddinott, 2000). Development programs with the aim of alleviating poverty in the long term should be especially concerned about these structurally poor households and thus require a poverty measure that allows their identification (Barrett, 2010). Hence, the *specificity* of expenditure-based indicators can be questioned.

Asset-based approaches to poverty measurement have been proposed to distinguish the structurally poor from the transient poor (Carter & Barrett, 2006). Besides being more cost-efficient and less demanding in terms of data requirements, this approach is robust to small fluctuations in poverty levels and, therefore, might be able to capture the structural component of poverty (Adato, Carter, & May, 2006; Barrett et al., 2006b). Several asset indices are already extensively used. A distinction can be made between those that use a theoretical and axiomatic framework and those that are primarily data-driven.

Of the indicators using a theoretical and axiomatic framework the Human Development Index (HDI) and the Multidimensional Poverty Index (MPI), both developed by the UNDP, are among the most frequently encountered (UNDP, 2010). The MPI is a weighted average of three pre-defined dimensions of poverty - education, health and standard of living- and can be decomposed into poverty headcount and intensity (Alkire & Santos, 2011). For each dimension several sub-indicators are included. A household is considered multidimensionally poor if it suffers deprivation in at least 33% of the weighted sub-indicators (Alkire & Santos, 2011). These indices were specifically developed to compare poverty between countries and over time and perform quite well in that context (Alkire & Santos, 2014). However, they might not be sufficiently sensitive to changes in poverty rates in specific local contexts, which casts doubt on their *relevance* for use in development programs. For such programs, the indicator developed by Zeller, Sharma, Henry, and Lapenu (2006) might be more useful. This indicator, like the MPI and HDI, has a theoretical basis but was developed to assess to which extent a policy or program reaches the poorest. A downside of this indicator is its reliance on principal component analysis (PCA), which makes it decidedly less easy-to-use and, thus, violates the *cost-effective Availability* criterion.[3]

The second group of poverty indicators does not use theoretical justification to select assets but selects assets solely on their statistical relationship with poverty. Because of their data-

[3] In the same vain is the asset index promoted by Sahn and Stifel (2000, 2003). They construct poverty-sensitive asset indices using factor analysis, a methodology comparable to PCA.

driven nature, these indicators are country-specific. Two frequently used indicators in this group are the Poverty Assessment Tool (PAT) and the Progress out of Poverty Index (PPI). They do not exist for all countries yet, but are under continuous development. The PAT has been developed by USAID (2014) and its use is mandatory for many USAID funded projects, while the PPI grew out of a microfinance initiative in Bosnia-Herzegovina (Matul & Kline, 2003) and was further developed by Mark Schreiner of Microfinance Risk Management L.L.C and the Grameen foundation. At the time of writing, the PPI was used by 176 organisations, mostly in the sphere of microfinance, including *Oiko Credit*, a financial cooperative supporting microfinance initiatives worldwide, and *Dia*, an organisation supporting MFIs in India (Grameen Foundation, 2014). Both tools are similar in many respects, although the PAT is slightly more accurate and the PPI more widely available (Schreiner, 2014). Neither of the indicators is frequently encountered in the academic literature, although the PPI has been used to assess program impact (Blauw & Franses, 2011; Larsen & Lilleor, 2013) and as a benchmark indicator (Dinh & Zeller, 2010).

As many organisations are interested in using PPI in program evaluation and impact, an independent evaluation of how SMART it is as an indicator for poverty, and the extent to which it can be used for targeting the poor and monitoring and evaluating development programs is highly relevant for policy makers and development professionals.

The PPI has been developed with the specific aim to measure poverty at household level in a particular country. Moreover, the tool has been designed to provide a cost effective and timely available proxy for poverty. Hence, the SMART principles *Specificity*, *Measurability*, *cost-effective Availability* and *Timely availability* are clearly met. The *Relevance* criterion, however, cannot be accurately assessed without validation. We distinguish two important, but different aspects of the *Relevance* criterion: the ability of the indicator to distinguish poor from non-poor households regardless of where they live and the sensitivity of the PPI to changes in poverty over time. The first point is important for targeting and reporting, while the second point is crucial for the indicator to be useful for monitoring and evaluation.

In this paper we focus on the PPI of Rwanda, which is particularly well-suited to assess the relevance of the PPI for two reasons. First, the PPI has been calibrated on a household survey conducted in 2010/2011 and a similar household survey has been conducted in 2010/2011. This ensures that we have perfectly comparable data. Second, in the 5-year interval between survey rounds the country has experienced considerable economic growth, creating changes in poverty rates (Ansoms & Rostagno, 2012). This is a necessary condition to assess the sensitivity of the PPI to changes in poverty rates over time.

In the next section we briefly explain how the PPI is constructed. Then, we outline the methodology and describe the data used to examine the validity of the PPI. In the results section we show that poverty estimates based on index scores corresponded well to official poverty rates and that the index was useful for reporting and targeting. Furthermore, we show that its sensitivity to changes in poverty over time depended crucially on a limited set of items; most items were stable, and did not change over time. In the final section the implications of these findings are discussed.

2 The Progress out of Poverty Indicator

PPIs have already been developed for more than 50 developing countries.[4] Their development is always based on detailed household-level data such as captured by the Living Standards

[4] www.progressoutofpoverty.org, accessed April 2014

Measurement Surveys of the World Bank or national household surveys and the methodology is standardized (Schreiner, 2010b). First, out of the household-level variables in the survey, a pre-selection of 100 indicators in the area of family composition, education, housing, and durable goods is made. Out of these, ten are selected that have a high correlation with poverty measured by the uncertainty coefficient (Goodman & Kruskal, 1979), are inexpensive to collect, easy to answer quickly, simple to verify, and liable to change over time as poverty status changes (Schreiner, 2010b). These ten items are given weights using logistic regression, such that final scores on the index range from 0 to 100. A scorecard is produced which allows users to calculate scores on the spot (figure 1). Using look-up tables, these scores can subsequently be converted into the likelihood that a household is below any one of a number of poverty lines (appendix A). In general tables are provided for 50%, 100% and 150% of the national poverty line, the food poverty line and an international poverty line such as the $1.25 (per person/day) line. Finally, the goodness-of-fit is assessed with out-of-sample calibration and standard errors for the likelihood of living below the poverty line given a PPI-score are obtained with bootstrapping. Country-specific details are provided in documentation available at the website of the Grameen Foundation (www.progressoutofpoverty.org).

The PPI scorecard for Rwanda was developed by Schreiner (2010b) following the methodology described above. It was calibrated on the national household survey EICV 2 (Enquête Intégrale sur les conditions de vie) conducted by the government of Rwanda in 2005/2006, which interviewed 6900 households. This survey was developed to monitor living conditions and covered all provinces of Rwanda. Data are publicly available from the website of the National Bureau of Statistics of Rwanda (GoR, 2006). The final PPI index for Rwanda (figure 1) contained two questions on household composition (Q1, Q4), two on education (Q2, Q3), four on housing conditions (Q5-Q8) and two on ownership of durable goods (Q9, Q10). Questions with higher discriminatory power in distinguishing poor from non-poor households were given a larger weight in the overall score. For instance, a maximum number of points (13) was attributed to households using candles, gas lamps, generators or the electrical grid as their main source of lighting (Q7). On the other hand, some responses only marginally increased the likelihood of living above the poverty line: households that sent all their children to school scored only two points more compared to households where at least one child had not gone to school in the last 12 months (Q2). By summing the points received for each question, a total score is obtained which can be converted into the likelihood the household is living below the poverty line using the provided look-up tables (appendix A). For instance, a household with a score between 50 and 55 had a likelihood of 22% of living below the national poverty line.

3 Data and methods

In this paper we use data from the 2010/11 EICV 3 survey to validate the 2005/2006 PPI scorecard of Rwanda. The EICV 3 expenditure survey contains 14,308 observations and used similar methodology and questionnaire as the earlier survey round, the EICV 2 (GoR, 2012), which was used by Schreiner (2010b) to develop the PPI for Rwanda. Both survey rounds used the same methodology and assumptions to construct the national poverty line and the food poverty line. These poverty lines were used to construct a poverty variable that indicates whether or not a household lived below the poverty or food poverty line, which is available for each household in the dataset.

In Rwanda, as is common for developing countries, poverty lines were calculated following the basic-needs approach (Martin Ravallion, 2012). This means that a basket of consumption goods corresponding with local dietary patterns that meets the minimum energy-

intake requirements was selected and converted to its monetary value. Based on the EICV 2, the food poverty line was set at 45,000 Rwandan Franc (RwF) per adult equivalent in 2001 prices.

Indicator	Value	Points	Score
1. How many household members are 17-years-old or less?	A. Five or more	0	
	B. Four	1	
	C. Three	7	
	D. Two	8	
	E. One	13	
	F. None	20	
2. Have all household members ages 7 to 17 been to school in the last 12 months?	A. No	0	
	B. Yes	2	
	C. No one in age range	3	
3. What is the highest grade that the female head/spouse has successfully completed?	A. Never attended school	0	
	B. Attended and completed none, one, or two years	2	
	C. Years 3 or 4 of primary	3	
	D. Years 5 or 6 of primary	5	
	E. There is no female head/spouse	5	
	F. Anything after 6 years of primary	9	
4. What is the status of the male head/spouse in his main occupation?	A. Agricultural wage worker, or does not work	0	
	B. There is no male head/spouse	3	
	C. Self-employed in agriculture, or unpaid worker (homemaker, apprentice, volunteer, etc.)	4	
	D. Non-agricultural wage worker	5	
	E. Self-employed in non-agriculture	8	
5. What is the main material of the floor?	A. Packed earth	0	
	B. Wood, cement, tiles, bricks, stone, or other	7	
6. How many rooms does the household occupy (do not count bathrooms, water closets, or kitchen)?	A. One	0	
	B. Two or three	5	
	C. Four	7	
	D. Five	9	
	E. Six or more	12	
7. What is the main source of lighting for the household?	A. Burning wood, or other	0	
	B. Home-made kerosene or fuel-oil lamp (*agatadowa*)	8	
	C. Candles, gas lamp, electrical grid, or generator	13	
8. What is the main fuel used for cooking?	A. Firewood, field waste, or other	0	
	B. Charcoal, LPG, electricity, or kerosene	16	
9. Does the household own a radio or radio-cassette player?	A. No	0	
	B. Yes	3	
10. How many ares of agricultural land does the household own or use?	A. 0 to 10	0	
	B. 11 to 35	1	
	C. 36 to 60	2	
	D. 61 to 100	4	
	E. 101 to 150	6	
	F. 151 or more	9	
Microfinance Risk Management, L.L.C., http://www.microfinance.com		**Total score:**	

Figure 1: Poverty scorecard of Rwanda developed by the Grameen Foundation
Source: Schreiner (2010)

The poverty line takes into account non-food expenditure on top of food expenditure and was set at 64,000 RwF per adult equivalent in 2001 prices.[5] These national poverty lines were adjusted for household composition, inflation, and difference in price levels between the Rwandan provinces to make them comparable over time and between regions. The methodology is described in detail in GoR (2012) and McKay and Greenwell (2007).

Because of the similarities between EICV 2 and EICV 3, the methodology to validate the PPI by comparing both surveys is rather intuitive and straightforward. First, the PPI score was

[5] In EICV 3, the national poverty and the food poverty line equalled 118000 RwF ($ 200) and 83,000 RwF ($140) per adult equivalent in 2011 prices, respectively

calculated for each household in the EICV 3 sample based on the scorecard calibrated on 2005/06 data (figure 1). This allows comparing the correlation between PPI scores and poverty rates and thus assess its *Relevance* as a proxy of poverty: a higher PPI score in 2010/11 needs to correspond with a lower probability of living below the poverty line and thus with a higher income. This relationship should hold for both rural and urban households of which poverty rates seem to have changed differently. Coding was straightforward[6] and only 13 out of 14,308 observations had to be dropped because of missing variables.

Second, the effectiveness of the PPI as a tool to report and target the poorest households included in a development program was assessed, i.e. assessing its *Relevance* in distinguishing poor from non-poor households. Given a cut-off value of the PPI for inclusion in a pro-poor development program, we determined the number of households that would correctly be identified as poor and the number of households that was non-poor but would nevertheless be included. The better the PPI is able to identify poor households and to exclude non-poor households, the higher its discriminatory power, and the smarter it is as poverty indicator.

Third, we checked the *relevance* of the PPI in tracking changes in poverty status over time. To this end, we compared the actual poverty rate in 2010/11 with the poverty rate predicted by the PPI converted into poverty rates with the tables calibrated and provided by Schreiner (appendix A). From these results, we determined the questions that contributed most to the overall change in the PPI over time.

4 Results

4.1 *Relevance:* distinguishing poor from non-poor households

To verify whether the PPI scorecard was still able to accurately identify poverty five years after having been developed, PPI scores were compared to actual poverty rates in 2010/11. A PPI score was calculated for all households in the EICV 3 dataset based on the scorecard for Rwanda (figure 1). The distribution of the PPI score was nearly normal, but slightly skewed towards the right (mean=41.85, sd=13.07). Almost 80% of households had a PPI score between 25 and 55, while none of the households had a score below 5 or above 95 (table 1).

Table 1 confirmed the internal validity of the PPI: a higher PPI score reduced the probability of being poor. For instance, a household with a PPI between 35 and 40 had a probability of 50% (23%) to live below the national (food) poverty line, while these probabilities decreased to 38% (15%) for households with PPI scores between 40 and 45. Similarly, the number of households belonging to the highest income quintile increased with higher PPI scores, while the number belonging to the lowest quintile decreased with higher scores.

4.2 *Relevance:* in both urban and rural areas

Ideally, urban and rural households with identical PPI scores would have the same probability of living below the poverty line. Table 2 shows the results for Rwanda. Given their PPI score, poverty rates were only slightly lower for urban than for rural households (see last column, table 2). For each range of PPI scores, the likelihood of being poor based on poverty lines was similar between rural and urban areas. The largest difference (of 13%) occurred for households with scores between 40 and 45.

[6] Only Q4 posed some problems as many households had several occupations, and it was difficult to clearly determine their main occupation. Hence, these households were always attributed the score that corresponded to the occupation with the highest weight. This assumption might lead to a slight underestimation of the poverty rates. As we will primarily be concerned with an overestimation of the poverty rates, this potential downward bias would not invalidate our results.

Table 1: Trends of PPI (calculated from 2010/11 data using the 2005/06 scorecard) by indicators of poverty in 2010/11

PPI	HH below poverty line (%)	HH below food poverty line (%)	HH in highest income quintile (%)	HH in lowest income quintile (%)	n
5-9	100.0	92.3	0.0	92.3	13
10-14	93.8	83.3	0.0	78.1	96
15-19	86.2	68.0	1.3	62.2	225
20-24	80.2	57.8	2.2	51.9	592
25-29	74.5	45.4	3.4	38.5	1219
30-34	61.3	33.3	4.5	28.2	2039
35-39	50.3	23.1	9.0	18.2	2556
40-44	37.8	15.0	13.5	10.9	2448
45-49	25.0	7.7	24.5	5.3	1795
50-54	14.0	3.7	36.4	2.8	1142
55-59	7.4	1.4	54.2	0.5	734
60-64	2.9	0.8	74.2	0.6	476
65-69	0.6	0.0	88.6	0.0	350
70-100	0.0	0.0	94.7	0.0	610

Source: Authors' calculations from the 2010/11 EICV 3

Note: the reported poverty line and food poverty line are the national poverty lines as calculated by the Government of Rwanda, set at 118,000 RwF and 83,000 RwF in 2011 prices, respectively.

The slight, but systematic, overestimation of urban poverty might be related to the question on land ownership (Q10). More than half of urban households indicated to own less than 10 ares of land, which does not necessarily signal poverty in an urban region. By including land in the index, there is an inherent underestimation of the expenditure of urban household. Hence, the validity and comparability of the PPI measure suffers from a trade-off between a poverty indicator that is valid country-wide and an indicator that is more accurate, but region-specific. However, the encountered overestimation was small, providing evidence that the bias in poverty estimates due to location between urban and rural areas should not be too problematic.

Table 2: Likelihood of living below the poverty line given a PPI score (calculated from 2010/11 data using the 2005/06 scorecard) for rural and urban households

PPI	Rural		Urban		Percentage point difference between rural and urban households
	HH below poverty line (%)	n	HH below poverty line (%)	n	
5-9	100.0	12	100.0	1	0
10-14	93.6	94	100.0	2	-6.4
15-19	85.8	212	92.3	13	-6.5
20-24	80.4	552	77.5	40	2.9
25-29	74.8	1121	70.4	98	4.4
30-34	61.3	1903	61.0	136	0.3
35-39	51.0	2370	41.9	186	9.0
40-44	38.9	2239	26.3	209	12.6
45-49	25.7	1626	17.8	169	8.0
50-54	14.7	965	10.2	177	4.5
55-59	8.8	532	3.5	202	5.4
60-64	3.9	231	2.0	245	1.9
65-69	0.8	122	0.4	228	0.4
70-100	0.0	169	0.0	441	0

Source: Authors' calculations from EICV 3

4.3 *Relevance:* reporting and targeting of poor households

The Grameen Foundation promotes the PPI as a reporting tool that allows quick identification of poor households included in a development program. In his paper describing the development of the scorecard for Rwanda, Schreiner (2010b) promotes it as a targeting tool, useful for selecting households to include in the program. The PPI does not directly identify whether or not a household is below a specific poverty line but reports likelihoods. To use the tool for targeting, a cut-off value has to be selected. Households with a score below the cut-off value are included, while households with a higher score are excluded from the program. Poverty likelihoods of individual households and poverty rates of groups of households can be determined through look-up tables provided with the scorecard (appendix A). These tables are constructed based on the same information as the scorecard itself, which in the case of Rwanda is the 2005/06 household survey round on which Schreiner (2010) calibrated the PPI scorecard.

For the sake of conciseness, we present results for a PPI cut-off value of 35; each household with a PPI below 35 would be included in our hypothetical pro-poor program, households with a PPI of more than 35 would be excluded. The 2005/2006 data (appendix B) predicts that 47% of households would be included, and the program would reach almost 70% of the poor Rwandese households (table 3). Suppose we start the program in 2011 and we use the same PPI cut-off value of 35 calculated using the same scorecard, but with household data collected in 2010/2011. As living conditions have improved, less households have a PPI score below 35 in 2010/11. The program now targets 29% of the Rwandese households, and includes 50% of the country's poor households. Hence while intending to reach 70% of the poor households (point A on figure 2), the program reaches only 50% of them (point B on figure 2). Figure 2 illustrates that the inclusiveness of poor households is missed by 20% when 'old' scorecards are used on 'new' data. Hence the scorecard seems not to be specific enough to capture poor groups.

Table 3: Comparison between estimated and actual targeting effectiveness with a cut-off values of PPI of 35 for a development programmes that starts in 2011

Cut-off value PPI of 35	% of HH targeted	% of poor who are targeted	% of targeted HH who are poor	Poor HH targeted per non-poor HH targeted
Estimated targeting effectiveness (calibration with 2005/06 data)[1]	47.4	68.1	77.4	3.4
Actual targeting effectiveness (calibration with 2010/11 data)[2]	29.3	50.3	70.0	2.3

[1] Source: Schreiner (2010a)(table 12, p75)
[2] Source: Authors' calculations from EICV 3

Moreover, the severity of mistargeting for different target poverty rates is illustrated in figure 2. The horizontal axis shows the proportion of poor households that the project initially aimed to include (which corresponds to different PPI cut-off values), while the vertical axis is the proportion of poor households that were actually included calculated on the 2010/11 data for each PPI cut-off value using the 'old' scorecard. Ideally, for each PPI cut-off value, the proportion of poor households targeted in our program would be equal using either 2005/06 or 2010/11 data (diagonal). Yet, proportions of poor households targeted for each PPI cut-off value calculated with 2010/11 data are substantially lower as shown by the curve below the diagonal which shows the relationship of the proportion of poor households targeted for each PPI cut-off value using 2005/06 data versus 2010/11 data. Projects starting in 2011 that based their cut-off values on

estimations on 2005/06 data would consistently reach fewer poor households than intended (the curve never crosses the diagonal).

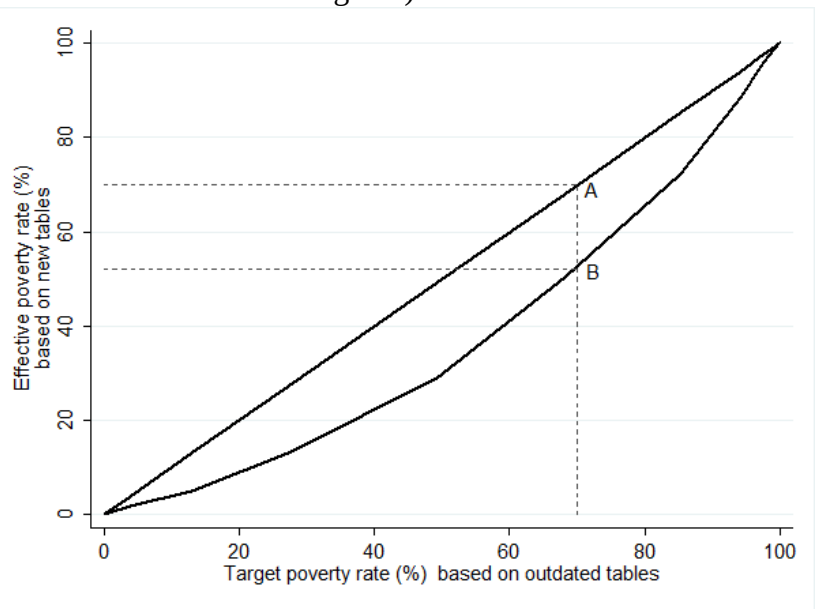

Figure 2: The number of effectively included poor households (vertical axis) in a development programme is always lower than initially intended (horizontal axis)
Source: Authors' calculations from EICV 3 and tables provided by Schreiner (2010a)(included in appendix, table 2)

Targeting effectiveness seems to be highly sensitive to the initial choice of PPI cut-off values. This effect was especially severe for projects choosing relatively low cut-off values, as the effect was most pronounced for cut-off values in the range of 30 to 50. As such, development projects only using the PPI as a reporting device for poverty inclusion would consistently overestimate the number of poor households that were actually reached by their project.

Another relevant question for development aid donors of pro-poor programs is whether the number of households that are non-poor who are nevertheless included in the development program is sensitive to an inaccurate choice of the cut-off value. For instance, based on the 2005/06 table (appendix B) a project with a cut-off value of 35 would have estimated that 77% of the targeted households were living below the poverty line. However, based on more recent data (table 3), only 70% of targeted households were living below the poverty line. On a positive note, this bias remained below 15 percentage points across all PPI cut-off values (results not shown but available upon request).

4.4 *Relevance:* monitoring program impact over time

To be a *Relevant* indicator for impact analysis, the PPI should capture changes in poverty rates over time. Schreiner (2010b) states explicitly that some questions such as radio ownership were purposely included in the Rwanda PPI scorecard because they were believed to be sensitive to changes in poverty rates over time. Yet, this was not formally tested by Schreiner (2010).

We test sensitivity to changes in poverty status by comparing estimated poverty rates based on the tables provided by Schreiner (and calibrated on the 2005/06 data) with actual poverty rates observed in 2010/11 based on the national poverty line. Poverty rates estimated with the outdated tables consistently overestimated actual poverty rates in 2010/11 (figure 3). However, overestimation was limited and never exceeded 10%. A similar analysis for food poverty lines confirmed these results (results not shown but available upon request).

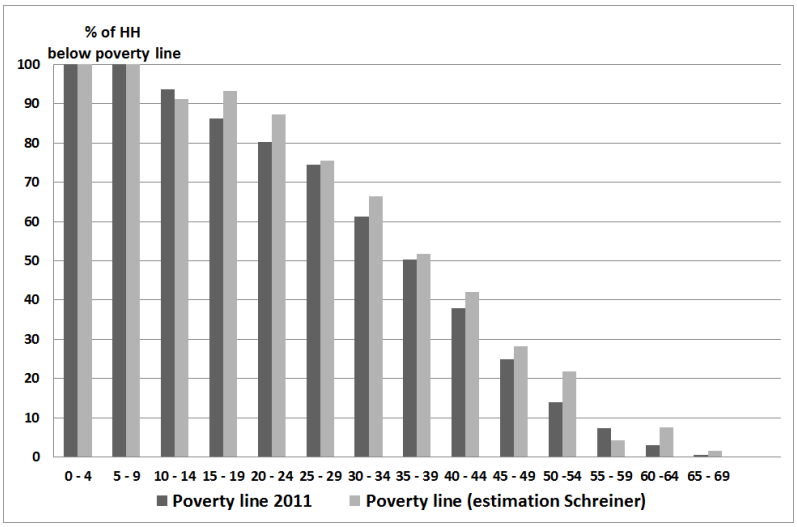

Figure 3: Comparison between actual number of HHs below poverty line in 2011 and the number estimated by Schreiner based on 2005 data

Source: Authors' calculations from EICV 3 and tables provided by Schreiner (included in appendix A)

Moreover, the poverty rates at provincial level in 2010/11 were also estimated based on Schreiner's tables (appendix A) and compared with the official poverty rates at provincial level (table 4). Besides a slight underestimation of poverty in the Southern province, the estimations based on the PPI were in line with government statistics. Hence, the PPI was sufficiently sensitive to monitor the decrease in poverty between 2005 and 2011. Therefore, at first glance it seems that the PPI achieves its objective of capturing structural improvements in poverty rates.

Table 4: Poverty rates in 2005/06 and 2010/11 estimated with PPI and compared with official government statistics

	2005/06		2010/11	
	PPI	**Official poverty rate**	**PPI**	**Official poverty rate**
Kigali City	22.0	20.8	16.1	16.8
Eastern Province	56.2	52.1	45.3	42.6
Northern Province	60.4	60.5	45.9	42.8
Southern Province	60.9	66.7	48.5	56.5
Western Province	60.9	60.4	47.1	48.4

Source: Authors' calculations and government report 'The evolution of poverty in Rwanda from 2000 to 2011: results from the household surveys (EICV) (GoR, 2012)'

A closer look at the ten indicators that constitute the PPI revealed that the improvement in the overall PPI score from an average of 37.2 in 2005 to 41.9 in 2011 was mainly driven by two indicators, namely main lighting source (Q7) and radio ownership (Q9).Together, these indicators contributed to an increase in the overall PPI score of 3.8 points, which accounted for 80% of overall change. Radio ownership increased from 14% in 2005 to 86% in 2011, contributing 1.5 points to the overall increase. This sharp increase in radio ownership was confirmed by other sources (Government of Rwanda, 2012). Note that it was also expected to happen by the developers of the scorecard. Similarly, the number of households achieving the maximum score of 13 on question 7[7], related to their main source of lighting, increased from 9% to 45%, which

[7] The exact question is phrased as follows: "What is the main source of lighting for the household?" A. Burning, wood or other (0 points); B. Home-made kerosene or fuel oil lamp (3 points); C. Candle, gas lamp, electrical grid or generator (13 points).

contributed 2.3 points to the overall change. This sharp increase was primarily explained by the fact that the highest attainable score for this item was attributed to households reporting in 2011 that battery-powered lanterns were their main source of lighting a response category which did not yet exist in 2005. As almost one third of the households reported using battery-powered lanterns, this effect is important. Consider the following: if households with access to battery-powered lanterns were classified in the lowest category, attributing a score of zero to this question, the PPI overestimated poverty by more than 10% for several ranges of PPI scores (results not shown). In this case, PPI would no longer be *Relevant* as an indicator to monitor changes in poverty rates over time.

Most of the other questions included in the PPI are rather insensitive to changes in poverty rates. Items such as household composition or arable land area per household are unlikely to change quickly as poverty decreases, such that almost all differences in poverty rates have to be picked up by only few items – such as radio ownership.

5 Discussion and conclusions

The objective of this paper was to assess the validity of the PPI, which was developed as an easy-to-use and quick-to-implement asset-based indicator, for use in development programs. Its validity was assessed using the SMART criteria as defined by the European Commission. By design, the PPI is certainly *Specific* as it serves a well-defined purpose, namely measuring poverty at household level in a specific region; *Measurable*, because the proxy quantifies the probability a is living household below the poverty line; *Available cost effectively*, because the indicator consists of only 10 easy questions; *Timely available*, because collecting the PPI can easily be administered at regular time intervals and the data can be processed quickly. Yet, concerns arise on its *Relevance* in distinguishing poor from non-poor household and in capturing changes of poverty over time. Its *relevance* was tested in this paper.

The relationship with expenditure poverty was analysed using a combination of the household survey round used to develop the PPI, the EICV 2, and its most recent version, the EICV 3. Overall, poverty estimates based on index scores were very close to officially reported poverty rates. The strength of this relationship was consistent between urban and rural areas, showing the robustness of the indicator to distinct living conditions within the country. The PPI was also correctly distinguishing poor from non-poor households, which makes it a useful targeting tool for development organisations.

Whether the PPI is accurate enough to use for reporting the number of poor households that were included in development programs depends crucially on which poverty thresholds and conversion tables were used. Our results indicate that a project starting in 2011 that would use the original conversion tables, which were developed five years earlier, might reach substantially fewer poor households than it had intended. This effect might be especially severe in the context of a country experiencing stark economic growth and substantial poverty reduction over the interval between scorecard development and field application, such as Rwanda (Ansoms & Rostagno, 2012). Consequently, reports of development projects might systematically overestimate the number of poor households participating in the program.

Whether reaching fewer poor household than initially intended or targeting more non-poor households poses a problem for a development program depends on the specific aim and context of the program. Moreover, it is never advisable to use one single index to decide on inclusion or exclusion of a household in a program (Skoufias, Davis, & De La Vega, 2001; Van de Walle, 1998) and it should be emphasized that the advocates of PPI do not advise to do so. Instead, they promote PPI as a reporting device. Although using PPI merely for reporting and not

targeting would not result in mistargeting, our results show that such reports would consistently overestimate the number of poor households reached by the program. Hence, the indicator fails on a number of accounts in terms of *relevance* in measuring poverty.

One of the reasons the PPI was developed was to evaluate project impact. Therefore, care was taken to include items in the scorecard that were likely to change with improved income over time. Although poverty rates were found to be consistently overestimated for almost all ranges of the index, the degree of overestimation did not exceed 10%, indicating the index indeed appears to be reasonably sensitive. However, this sensitivity depended crucially on a limited set of items, with two items being responsible for over 80% of the observed variation and most items showing no significant change over time at all. This drawback is recognized by the Grameen foundation because poverty scorecards are updated as soon as new country-wide expenditure surveys become available. However, in a region with a sharp decrease in poverty rates over a short time-span such as Rwanda, an update of the PPI every 5 years might be insufficient to guarantee effective targeting of the poorest households and sensitivity to poverty status changes to evaluate the impact of development projects. Moreover, as soon as a PPI is updated, the baseline study will probably be outdated because this study will not necessarily include all the items of the new PPI. Consequently, it would no longer be possible or, at least, questionable to measure program impact over time by combining the old and new indicator. Perhaps the general trend in poverty dynamics could still be assessed through estimating the poverty rates in the baseline year with the old PPI and estimating poverty in subsequent years with the most recent PPI. An additional disadvantage of the need to update the PPI regularly are the costs involved in designing and disseminating the new PPI.

Another concern is that the indicators might be biased upwards by construction. Households that escape poverty might indeed buy a radio or increase the number of rooms in their dwelling, but households confronted with a negative shock which pushes them back into poverty are less likely to sell their radio or decrease the number of rooms in their house. Although this seems intuitive, without panel data with sufficient variation this assertion could not be tested directly. Furthermore, given land tenure systems prevalent in developing countries like Rwanda, it is unlikely that land will be sold. Given that the indicator was developed to be sensitive to changes in poverty, its sensitivity to negative shocks is an important consideration. Such sensitivity to both up- and downward movements in poverty is also crucial to make it a valuable indicator to study poverty dynamics and poverty traps (Carter & Barrett, 2006).

In conclusion, for such a relatively simple and easy-to-use indicator, the PPI does a remarkable job in estimating poverty levels and can be considered a SMART indicator, with some reservation regarding the *Relevance* component within the SMART framework. It is *Relevant* for development organisations to provide a description of project participants. It is also useful for targeting, although its accuracy might be compromised when the scorecard is several years old. However, its sensitivity to changes in poverty is limited and warrants further study. Hence, some hesitation is required before using the index as a tool to evaluate the impact of development projects.

References

Adato, M., Carter, M. R., & May, J. (2006). Exploring poverty traps and social exclusion in South Africa using qualitative and quantitative data. *Journal of Development Studies, 42*(2), 226-247. doi: 10.1080/00220380500405345

Alkire, S., & Santos, M. E. (2011). Acute multidimensional poverty: A new index for developing countries. *Proceedings of the German Development Economics Conference, Berlin 2011, No. 3*.

Alkire, S., & Santos, M. E. (2014). Measuring Acute Poverty in the Developing World: Robustness and Scope of the Multidimensional Poverty Index. *World Development, 59*(0), 251-274. doi: http://dx.doi.org/10.1016/j.worlddev.2014.01.026

Ansoms, A., & Rostagno, D. (2012). Rwanda's Vision 2020 halfway through: what the eye does not see. *Review of African Political Economy, 39*(133), 427-450.

Barrett, C. B. (2010). Measuring Food Insecurity. *Science, 327*(5967), 825-828. doi: 10.1126/science.1182768

Barrett, C. B., Marenya, P. P., McPeak, J., Minten, B., Murithi, F., Oluoch-Kosura, W., . . . Wangila, J. (2006a). Welfare dynamics in rural Kenya and Madagascar. *The Journal of Development Studies, 42*(2), 248-277.

Barrett, C. B., Marenya, P. P., McPeak, J., Minten, B., Murithi, F., Oluoch-Kosura, W., . . . Wangila, J. (2006b). Welfare dynamics in rural Kenya and Madagascar. *Journal of Development Studies, 42*(2), 248-277. doi: 10.1080/00220380500405394

Baulch, B., & Hoddinott, J. (2000). Economic mobility and poverty dynamics in developing countries. *The Journal of Development Studies, 36*(6), 1-24. doi: 10.1080/00220380008422652

Beegle, K., De Weerdt, J., Friedman, J., & Gibson, J. (2012). Methods of household consumption measurement through surveys: Experimental results from Tanzania. *Journal of Development Economics, 98*(1), 3-18.

Blauw, S. L., & Franses, P. H. (2011). *The impact of mobile telephone use on economic development of households in Uganda*. Tinbergen Institute Discussion Paper. Tinbergen Institute.

Carter, M. R., & Barrett, C. B. (2006). The economics of poverty traps and persistent poverty: An asset-based approach. *Journal of Development Studies, 42*(2), 178-199. doi: 10.1080/00220380500405261

Deaton, A. S. (1997). *The Analysis of Household Surveys: Microeconometric Analysis for Development Policy*. Baltimore: Md. John Hopkins University Press for the World Bank.

Dinh, T. T. V., & Zeller, M. (2010). *Development of operational poverty indicators in Northern Vietnam*. Paper presented at the Sustainable land use and rural development in mountainous regions of Southeast Asia, Hanoi.

European Evaluation Network for Rural Development. (2014). *Defining proxy indicators for rural development programmes (working paper)*.

Goodman, L. A., & Kruskal, W. H. (1979). *Measures of association for cross classification*. New York, NY: Springer-Verlag.

GoR. (2006). *Preliminary poverty update report: integrated living conditions survey 2005/06*. Kigali: National Institute of Statistics of Rwanda.

GoR. (2012). *The evolution of poverty in Rwanda from 2000 to 2011: results from the household surveys (EICV)*. Kigali: National Institute of Statistics of Rwanda.

Government of Rwanda. (2012). Rwanda Demographic and Health Survey 2010: Final Report. Kigali, Rwanda: National Institute of Statistics of Rwanda: Ministry of Finance and Economic Planning.

Grameen Foundation. (2014). Progress out of poverty Retrieved March 10, 2014, from http://www.progressoutofpoverty.org/

Larsen, A. F., & Lilleor, H. B. (2013). *Beyond the field: Impact of farmer field schools on food security and poverty alleviation*. Copenhagen, Denmark: The Rockwool Foundation Research Unit.

Matul, M., & Kline, S. (2003). Scoring change: Prizma's approach to assessing poverty: Microfinance Centre for Central and Eastern Europe and the New Independent States.

McKay, A., & Greenwell, G. (2007). Methods used for poverty analysis in Rwanda update note. In O. P. Management (Ed.).

Poister, T. H. (2008). *Measuring performance in public and nonprofit organizations*: John Wiley & Sons.

Ravallion, M. (1996). *Issues in measuing and modeling poverty*. Washington, USA: World Bank.

Ravallion, M. (2012). *Poverty lines across the world*. Oxford Handbook of the Economics of Poverty.
Ravallion, M., Datt, G., & Walle, D. (1991). Quantifying absolute poverty in the developing world. *Review of income and wealth, 37*(4), 345-361.
Sahn, D. E., & Stifel, D. (2000). Poverty Comparisons Over Time and Across Countries in Africa. *World Development, 28*(1), 2123-2155.
Sahn, D. E., & Stifel, D. (2003). Exploring Alternative Measures of Welfare in the Absence of Expenditure Data. *Review of Income and Wealth, 49*(4), 463-489. doi: 10.1111/j.0034-6586.2003.00100.x
Schreiner, M. (2010a). Seven extremely simple poverty scorecards. *Enterprise Development and Microfinance, 21*(2), 118-136.
Schreiner, M. (2010b). A Simple Poverty Scorecard for Rwanda.
Schreiner, M. (2014). How do the poverty scorecard and the PAT differ? microfinance.com.
Skoufias, E., Davis, B., & De La Vega, S. (2001). Targeting the poor in Mexico: an evaluation of the selection of households into PROGRESA. *World Development, 29*(10), 1769-1784.
UNDP. (2010). The real wealth of nations: Pathways to human development *Human Development Report*. New York, USA: United Nations Development Programme.
USAID. (2014). Poverty Assessment Tools Retrieved March 10, 2014, from http://www.povertytools.org/
Van de Walle, D. (1998). Targeting revisited. *The World Bank Research Observer, 13*(2), 231-248.
Zeller, M., Sharma, M., Henry, C., & Lapenu, C. (2006). An operational method for assessing 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.

Appendices

Appendix A: Estimation of likelihood of living in poverty given a PPI score (calibrated on EICV 2, provided by Schreiner (2010))

PPI	% below poverty line	% below food poverty line
0-4	100.0%	100.0%
5-9	100.0%	100.0%
10-14	91.2%	85.4%
15-19	93.2%	81.3%
20-24	87.3%	65.5%
25-29	75.5%	52.4%
30-34	66.5%	42.7%
35-39	51.8%	27.1%
40-44	42.1%	15.3%
45-49	28.1%	5.7%
50-54	21.8%	12.0%
55-59	4.2%	1.5%
60-64	7.5%	4.3%
65-69	1.6%	0.0%
70-74	0.0%	0.0%
75-79	0.0%	0.0%
80-84	0.0%	0.0%
85-89	0.0%	0.0%
90-94	0.0%	0.0%

Source: Schreiner (2010a)(table 4, p66 and table 4; p77)

Appendix B: Effectiveness of PPI in targeting the poorest household for different cut-off values of PPI in Schreiner (2010)

Cut-off value	% of HH targeted	% of poor who are targeted	% targeted who are poor	Poor targeted versus non-poor targeted
5	0	0	100.0	Only poor targeted
10	0.4	0.7	100.0	Only poor targeted
15	2.3	3.9	91.9	11.3
20	7.6	13	92.0	11.6
25	16.3	27.3	90.1	9.1
30	31.3	49.3	84.9	5.6
35	47.4	68.1	77.4	3.4
40	64.7	85.3	71.0	2.4
45	76.2	94.1	66.5	2.0
50	84.3	97.7	62.4	1.7
55	89.2	99.3	59.8	1.5
60	92.4	99.6	58.0	1.4
65	94.6	99.9	56.9	1.3
70	96.9	100	55.5	1.2
75	98.3	100	54.7	1.2
80	99.3	100	54.2	1.2
85	99.8	100	54.0	1.2
90	99.9	100	53.9	1.2

Source: Schreiner (2010a) (table 12, p75)

Appendix C: Effectiveness of PPI in targeting the poorest household for different cut-off values of PPI calibrated on 2010/11 data

Cut-off value PPI	% of HH targeted	% of poor[1] who are targeted	% of targeted HH who are poor	Poor HH targeted per non-poor HH targeted
10	0.1	0.2	100.0	only poor HH targeted
15	0.8	1.8	94.5	17.2
20	2.3	5.1	88.9	8.0
25	6.5	13.3	83.4	5.0
30	15.0	28.9	78.3	3.6
35	29.3	50.3	70.0	2.3
40	47.1	72.4	62.6	1.7
45	64.3	88.4	56.0	1.3
50	76.8	96.0	50.9	1.0
55	84.8	98.8	47.4	0.9
60	90.0	99.7	45.1	0.8
≥65	≥93.3	100	≤43.6	≤0.8

Source: Authors' calculations from EICV 3, but methodology based on Schreiner (2010a)(figure 12, p75).