

A Simple Poverty Scorecard for Haiti

Mark Schreiner

October 6, 2006

Microfinance Risk Management, L.L.C.
6970 Chippewa St. #1W, Saint Louis, MO 63109-3060, U.S.A.
Telephone: +1 (314) 481-9788, <http://www.microfinance.com>

Abstract

How poor are participants in development projects in Haiti? Based on data from a national survey, this paper presents an easy-to-use, objective poverty scorecard that uses 10 simple indicators to estimate the likelihood that a participant lives on less than \$1/day. With 90-percent confidence, estimated poverty rates are correct within ± 1.9 percentage points. With the low-tech scorecard, field workers can compute scores by hand in real time, helping programs target services, track changes in poverty over time, and report on poverty rates.

Acknowledgements

This work was funded by Grameen Foundation U.S.A. I am grateful FAFO and Willy Egset for the data and to Nigel Biggar, Rob Fuller, and Jeff Toohig. Anne Hastings and Julie Hubble of Fonkoze provided the translated scorecard.

A Simple Poverty Scorecard for Haiti

1. Introduction

This paper presents an easy-to-use, objective poverty scorecard to help development projects in Haiti to target services, track changes in poverty over time, and report on participants' poverty rates.

The low-tech scorecard is derived from the 2001 *Enquête sur les Conditions de Vie en Haïti* (ECVH) by the Institut Haïtien de Statistique et d'Informatique in collaboration with the Norwegian institute FAFO. Indicators were selected to be:

- Inexpensive to collect, easy to answer, and simple to verify
- Strongly correlated with poverty
- Liable to change as poverty status changes over time

The 10-indicator scorecard applies to both rural and urban Haiti. All scorecard weights are non-negative integers. Scores range from 0 (most-likely poor) to 100 (least-likely poor). Field workers can compute scores by hand in real time.

A participant's score corresponds to a "poverty likelihood", that is, the probability of being poor. The share of all participants who are poor is the average poverty likelihood. For participants over time, progress is the change in their average poverty likelihood.

The scorecard is accurate in that poor people are concentrated among low scores and non-poor people among high scores. Furthermore, the estimated poverty likelihoods of individuals and the overall poverty rate of groups are quite close to their true values.

Precision was measured by bootstrapping hold-out samples. The 90-percent confidence intervals for estimated poverty likelihoods are about ± 6.5 percentage points, and the 90-percent interval for estimated overall poverty rates is ± 1.9 percentage points.

The scorecard is an appropriate tool for USAID microenterprise grantees who are required to report (in a proven, objective way) the share of participants who live on less than \$1/day.

Grameen Foundation U.S.A. can use the scorecard to classify participants as very poor (poorest half in poverty), moderately poor (top half in poverty), and non-poor.

2. A poverty line for Haiti

There is no official poverty line for Haiti. Based on a subsistence level of food plus basic non-food consumption, Pedersen and Lockwood (2001) use the 1999 *Enquête Budget et Consommation des Ménages* (EBCM) to estimate an expenditure-based poverty line of 5,638 Gourdes/person/year. This poverty line reflects international best practice. Unfortunately, it could not be used in this paper because the EBCM lacks data on most indicators typically found in poverty scorecards.

Instead, this paper uses the 2001 ECVH, which lacks expenditure but has most key poverty indicators. The poverty line is income-based.

Sletten and Egset (2004) compute a \$1/person/day (purchase-power parity) income-based poverty line of 2,757 Gourdes/person/year. This does not adjust for regional cost-of-living (Haiti does not have sub-national price indices), but it does facilitate international comparisons. By this line and using the 2001 ECVH, 55.5 percent of Haitians were poor.

The EBCM's expenditure-based poverty line of 5,638 Gourdes should not be compared to ECVH's income-based poverty line of 2,757 Gourdes. In particular, it appears that implicit consumption from goods (in particular, food) produced by a household for its own consumption was undercounted in the ECVH. Furthermore, the EBCM line is based on what Haitian household actually consume, while the ECVH line is based—for lack of a better alternative—on an arbitrary international benchmark. Still, the two lines do not give wildly different poverty rates: 48 percent in 1999 for the EBCM (Pedersen and Lockwood, 2001) versus 55.5 percent for the EVCH in 2001.

This paper, of necessity and to fit USAID requirements, uses the \$1/person/day line. By this measure, Haiti is the poorest country in the Western Hemisphere, as well as the most unequal (Sletten and Egset, 2004).

3. Other poverty scorecards for Haiti

When the work reported here was almost complete, it was revealed that Fuller (2006) had already created a poverty scorecard for Fonkoze, the Haitian affiliate of Grameen Foundation U.S.A. That scorecard is similar in most ways to the scorecard here, as it:

- Uses the same data from the 2001 ECVH
- Uses the same methods as in Schreiner *et al.* (2004) and Schreiner (2006a-e)

Furthermore, Fuller’s scorecard is almost as accurate as the one here. The main differences between the scorecards are:

- Fuller uses only households with an adult female
- Fuller uses only indicators already collected by Fonkoze

An advantage of Fuller’s scorecard is that Fonkoze’s managers know Fuller and so are less likely to view the scorecard as a magic box of unknown origin, increasing the likelihood that they will use the scorecard. Of course, this advantage may not carry over to other Haitian development organizations.

4. Indicator preparation

About 250 potential poverty indicators were prepared for the 7,168 households in the 2001 ECVH. Broadly, the indicators cover:

- Household demographics (such as school attendance or number of children)
- Characteristics of the homestead (type of wall, type of toilet, or hectares of land)
- Household consumption (such as milk or meat)
- Household durable goods (such as televisions or machetes)

As a first step, the ability of each indicator to predict poverty was tested with the “uncertainty coefficient”, an entropy-based measure (Goodman and Kruskal, 1979).

About 100 indicators were selected for further analysis. These are listed in Figure 1 and ranked by the uncertainty coefficient. They are worded as in the English translation of the ECVH by FAFO, with possible responses ordered starting with those most strongly linked with poverty.

Some indicators in Figure 1 have similar relationships with poverty. For example, households with electricity are also more likely than other households to own a television. If a scorecard includes “owns a television”, then also including “electrical connection” adds little. In such cases, the scorecard uses only one of the indicators.

The scorecard also aims to measure *changes* in poverty through time. Thus, some indicators that are unlikely to change even if poverty changes (such as the highest grade completed by the female spouse/head) are omitted in favor of indicators that are less-powerful but more likely to change.

Some potential indicators were not selected for Figure 1 or for the scorecard because they are difficult to collect (“Did anyone eat any vegetables yesterday?”), difficult to compute (“What is the dependency ratio of children to adults?”), or too subjective (“If you had to borrow 500 Gourdes in the next seven days, could you?”). Also, most indicators related to past consumption were omitted, as field agents cannot straightforwardly verify whether the respondent’s memory or motives are trustworthy for questions such as “Did anyone in the household drink alcohol yesterday?”

Some indicators (such as “Is the household connected to the electrical grid?” and “Did the household eat rice yesterday?”) were omitted because the share of “Yes” respondents in the survey strained credulity (Fuller, 2006). Additional indicators (such as “In what type of dwelling does the household reside?” and “Did the household use solar energy for light?”) were omitted because it is unclear what the responses in the ECVH mean.

Finally, some powerful indicators were not used because managers and focus groups at Fonkoze suggested that the questions cause shame and thus were unlikely to elicit truthful answers. Examples include “How many meals were prepared in the household yesterday” and “Are all or any of the rooms in the residence, including the hall or kitchen, leaking in the roof?”.

In sum, scorecard indicators were selected both for the strength of their correlation with poverty and for the feasibility of collecting them easily and accurately.

Figure 1: Poverty indicators ranked by their uncertainty coefficient

<u>Uncertainty coefficient</u>	<u>Indicator (Responses ordered starting from the one most-closely linked with poverty)</u>
84	If your household had a sudden need for 500 Gourdes, would you be able to raise the money within a week? (No, it would be impossible; Perhaps, but I doubt it, or We would get some help from others; We would use our savings)
82	Does anyone in your household own a television set? (No; Yes)
82	Where does the household reside? (Not Port-a-Prince; Port-a-Prince)
79	If your household had a sudden need for 250 Gourdes, would you be able to raise the money within a week? (No, it would be impossible; Perhaps, but I doubt it, or We would get some help from others; We would use our savings)
74	Does anyone in your household own an electric fan? (No; Yes)
72	In what type of dwelling does the household reside? (<i>Kay até</i> (roof and walls merged); Hovel (<i>taudis/ajoupa</i>); Ordinary 1-level dwelling; Ordinary 2-or-more-level dwelling, apartment, villa, or colonial-type dwelling)
68	Does the residence have electricity? (No; Yes)
67	What is the main construction material of the floor of the residence? (Earth; Concrete or other; Ceramic or wood planks)
66	What is the highest educational level attained by any household member? (None to Basic 5; Basic 6 to Basic 10; Seconde, Rheto, or Philo; Technical professional cycle, post-secondary, or post-graduate)
66	Yesterday, did anyone in the household drink any alcoholic beverages, including liquors, syrups, clairin, cremas, beer, wine, rum, or other? (No; Yes)
64	Does anyone in your household own a clock? (No; Yes)
63	Does the household own any bulls, cows, horses, donkeys, or mules? (No; Yes)
63	Does the household own any horses, donkeys, or mules? (No; Yes)
63	What is the principle source of daily-use water for the household? (Rain, river, spring, other, or unknown; Public fountain or a well on the plot or in the surroundings; Tank truck, water brought by bucket, or bottled water; Water piped into the living quarters or into the compound)
62	What is the main construction material of the roof of the residence? (Straw, palm leaves, or other; Iron; Concrete)
62	Can the female head/spouse read a letter or newspaper and write a letter in French? (No; Can read and write, but not easily; Can read and write easily)
62	Does anyone in the household own a mix-master/electric blender? (No; Yes)
62	Does anyone in the household own a radio/cassette player? (No; Yes)
61	Does the household own any pigs? (No; Yes)
61	Does the household own any bulls? (No; Yes)
60	Does the household own any bulls or cows? (No; Yes)
60	What is the principle source of drinking water for the household? (Rain, spring, river, or other; Public fountain or a well on the plot or in the surroundings; Tank truck, water brought by bucket, or bottled water; Water piped into the living quarters or into the compound)
60	Does the household sharecrop-in any land? (No; Yes)
60	Does the household own a floor fan? (No; Yes)

Figure 1: Poverty indicators ranked by their uncertainty coefficient (cont.)

<u>Uncertainty coefficient</u>	<u>Indicator (Responses ordered starting from the one most-closely linked with poverty)</u>
60	Does the household own any goats? (No; Yes)
60	Does anyone in the household have a savings account at a bank or other formal financial institution? (No; Yes)
59	Does the household own a pick? (No; Yes)
59	Does the household own a hoe? (No; Yes)
58	Does the household own agricultural land? (No; Yes)
58	Does the household own a billhook? (No; Yes)
58	Does the household own a machete? (No; Yes)
58	Does the household own any cows? (No; Yes)
58	Does the household own a machete or a billhook? (No; Yes)
58	Does the household own, rent, or otherwise have access to any agricultural land? (No; Yes)
58	If you wanted to or needed to eat meat, chicken, or fish at least three times per week, could your household afford it? (No; Yes)
57	Is the household located in an urban area? (No; Yes)
56	Does the household own a refrigerator? (No; Yes)
51	In the past month, how often had you had to purchase food on credit because of lack of food or money to buy food? (Once a week or more; Sometimes, but less than once a week; Never)
53	How many agricultural plots does the household have? (1; 2 or 3; 4 or more; None)
51	What is the highest grade completed by the male head/spouse? (None to Basic 1; Basic 2 to Basic 5; Basic 6 to Seconde; Rheto or higher)
51	In the past month, how often had you had to limit the intake of adults to ensure that children get more because of lack of food or money to buy food? (Three times per week or more; Sometimes, less than three time per week; Never)
48	Does the household own a stove (electric, propane, or kerosene)? (No; Yes)
48	If you wanted to or needed to buy new clothes rather than second-hand clothes, could your household afford it? (No; Yes)
47	Can the male head/spouse read a letter or newspaper and write a letter in French? (No; Can read and write, but not easily; Can read and write easily)
46	What is the highest grade completed by the female head/spouse? (None to Basic 2; Basic 3 to Basic 6; Basic 7 to Basic 9; Basic 10 or higher)
46	How many meals were prepared in the household yesterday? (0 or 1; 2; 3 or more)
44	How many people from ages 0 to 14 live in the household? (4 or more; 3; 2; 1; 0)
44	Does the residence have toilet facilities? (No; Yes)
44	What is the main construction material of the walls of the residence? (Earth or other; wood or plywood; Concrete, bricks, blocks, or stones)

Figure 1: Poverty indicators ranked by their uncertainty coefficient (cont.)

<u>Uncertainty coefficient</u>	<u>Indicator (Responses ordered starting from the one most-closely linked with poverty)</u>
43	How many people from ages 0 to 17 live in the household? (5 or more; 3 or 4; 2; 1; 0)
43	Do any household members have salaried employment outside of the family business? (No; Yes)
43	How many people from ages 0 to 11 live in the household? (4 or more; 3; 2; 1; 0)
42	Where does the household usually store garbage before getting rid of it? (Nowhere or not in a container; In a container)
38	Can the female head/spouse read a letter or newspaper and write a letter in her mother tongue? (No; Can read and write, but not easily; Can read and write easily)
37	Do all children ages 6 to 14 attend school? (No; Yes; No children in this age range)
37	In the past 12 months, did any member of the household receive money or in-kind transfers from anyone abroad? (No; Yes)
36	Do all children ages 6 to 17 attend school? (No; Yes; No children in this age range)
36	What is the ratio of children 17 or younger to adults 18 or older? (≥ 0.5 ; <0.5)
35	Does the household own a video player? (No; Yes)
33	Does the household have a bathtub or shower, be it private or shared with neighbors? (No; Yes)
33	How many people from ages 0 to 5 live in the household? (3 or more; 2; 1; 0)
33	Do all children ages 6 to 11 attend school? (No; Yes; No children in this age range)
30	How many people live in the household? (7 or more; 5 or 6; 4; 3; 2 or 1)
30	Does the household own an oven (electric, propane, or kerosene)? (No; Yes)
29	How does the household store water? (Buckets or plastic gallon jugs; Drums or water tanks)
28	Do any close relatives (parents, spouses, children, or siblings of household members) age 14 or older live abroad? (No; Yes)
27	Can the male head/spouse read a letter or newspaper and write a letter in his mother tongue? (No; Can read and write, but not easily; Can read and write easily)
25	Do all male children ages 6 to 17 attend school? (No; Yes; No male children in this age range)
25	Is the household supplied with water by a private or public water company or institution? (No; Yes)
25	In the past month, how often had you had to reduce the number of meals eaten in a day because of lack of food or money to buy food? (Three times per week or more; Once or twice per week; Sometimes, but less than once per week; Never)
25	Are all or any of the rooms in the residence, including the hall or kitchen, leaking in the roof? (No; Yes)
23	Do all male children ages 6 to 14 attend school? (No; Yes; No male children in this age range)
22	Yesterday, did anyone in your household eat any fats or oils, including margarine, cooking butter, soya-oil, maize-oil, olive-oil, animal fat, or manteque? (No; Yes)
21	What type of toilet facility does the household use? (None; Private toilet or toilet shared with neighbors, or a hole on the plot or compound; Modern water closet)
21	Do all female children ages 6 to 14 attend school? (No; Yes; No female children in this age range)
21	Does any household member attend a school taught primarily in French? (No; Yes)
21	What is the employment status of the male head/spouse? (Self-employed, not employed, or no male head/spouse; All others)
21	Do you own or rent your residence and the plot on which it stands? (Own; Rent)

Figure 1: Poverty indicators ranked by their uncertainty coefficient (cont.)

<u>Uncertainty coefficient</u>	<u>Indicator (Responses ordered starting from the one most-closely linked with poverty)</u>
20	Is there a road accessible for cars leading to your dwelling? (No road; Dirt, gravel, or partly-paved road; Paved road)
20	Do all female children ages 6 to 11 attend school? (No; Yes; No female children in this age range)
19	Did anyone in the household have income from the sale of agricultural crops and products? (Yes; No)
19	Do all female children ages 6 to 17 attend school? (No; Yes; No female children in this age range)
18	Are all or any of the rooms in the residence, including the hall or kitchen, invaded by rats or mice? (No; Yes)
17	In the past month, how often have you had to go entire days without eating because of lack of food or money to buy food? (Sometimes; Never)
16	Do all male children ages 6 to 11 attend school? (No; Yes; No male children in this age range)
12	Does anyone in your household own a sewing machine? (No; Yes)
10	Does anyone in your household own a bed (mattress, box, iron-made)? (No; Yes)
5	What is the ratio of men to women, ages 18 and older? (<0.4 ; $=>0.4$)
0.02	In the past 12 months, has any credit been obtained outside the household for self-employment activities? (No; Yes)

5. Selecting indicators for the scorecards

An appropriate statistical approach for classifying people as poor/non-poor is Logit regression. Indicators were selected by combining statistics with the judgment of an analyst with expertise in both scoring and microfinance:

1. Start with a scorecard with no indicators
2. For each candidate indicator not already in the scorecard:
 - A. Add the indicator to the scorecard
 - B. Derive weights with Logit
 - C. Record the improvement in general accuracy measured by the “c” statistic
3. Select an indicator based on (Schreiner *et al.*, 2005; Zeller, 2004):
 - A. Likelihood of acceptance by users:
 - i. “Face validity” (experience, theory, and common sense)
 - ii. Simplicity and cost of collection
 - B. Likelihood of changing as poverty status changes
 - C. Accuracy
 - D. Contrast with indicators already in the scorecard
 - E. Verifiability and susceptibility to strategic falsification
4. Add the selected indicator to the scorecard
5. Repeat steps 2–4 until there are 10 indicators
6. Transform the original Logit weights so that:
 - A. All weights are non-negative integers
 - B. The minimum score is 0 (most likely poor), and the maximum is 100

This “MAXC” algorithm for Logit is analogous to the “MAXR” algorithm for ordinary-least squares in Zeller, Alcaraz V., and Johannsen (2005 and 2004); Zeller and Alcaraz V. (2005a and 2005b); and IRIS (2005a and 2005b). If all classification errors are equally costly, then R^2 and “c” are good general measures of accuracy.

“c” is the area under a Receiver Operator Characteristic curve (Baulch, 2003) that plots the share of poor people (vertical axis) versus the share of all people ranked by score (horizontal axis). It can also be seen as the share of all possible pairs of poor and non-poor households in which the poor household has a lower score. Finally, it is equivalent to the Mann-Whitney U statistic.

Figure 2: Haiti poverty scorecard in English

Indicator	Value	Points	Total
1. How many people in the household are 14 years old or younger?	Four or more	0	
	Three	3	
	Two	8	
	One	11	
	None	19	
2. Do all children of ages 6–14 attend school?	No	0	
	Yes	3	
	No children ages 6–14	3	
3. Where does the household reside?	Not Port-a-Prince	0	
	Port-a-Prince	15	
4. Does the household own a radio/cassette player?	No	0	
	Yes	7	
5. What are the dwelling's floors made of?	Earth	0	
	Concrete or other	4	
	Ceramic or wood planks	12	
6. In the past 12 months, did the household receive any money or gifts remitted from abroad?	No	0	
	Yes	7	
7. Does any household member have salaried employment?	No	0	
	Yes	12	
8. How many plots of agricultural land, forest land, pasture land, or gardens does the household use?	None	2	
	One	0	
	Two or three	5	
	Four or more	11	
9. What is the dwelling's roof made of?	Straw, palm leaves, or other	0	
	Iron	4	
	Concrete	9	
10. Does the household own any pigs?	No	0	
	Yes	5	
		Total:	

Figure 2: Haiti Kat Evalyasyon

		Pwen
1. Konbyen moun nan kay la ki genyen katòzan (14) ou byen ki gen mwens pase katòzan (14)?	4+	0
	3	3
	2	8
	1	11
	0	19
2. Eske tout timoun ki gen laj soti sis (6) lane rive katòz lane (14) al lekòl?	Pa tout timoun (6-14) al lekòl	0
	Wi	3
	Pa gen timoun (6-14)	3
3. Ki kote ou rete?	Deyò Pòtoprens	0
	Pòtoprens	15
4. Eske ou gen radyo oubyen radyo-kasèt nan kay la?	Non	0
	Wi	7
5. Koman atè andan kay ou fèt?	Tè	0
	Siman ou lòt	4
	Seramik ou planche	12
6. Eske nan douz mwa ki sot pase la fanmi an resevwa lajan ki soti nan peyi lòtbò dlo?	Non	0
	Wi	7
7. Eske gen moun nan fanmi an kap travay (salariaje)?	Non	0
	Wi	12
8. Konbyen mòso Tè ou Konbyen jaden fanmi an itilize?	0	2
	1	0
	2-3	5
	4+	11
9. Ak ki sa tèt kay la kouvri?	Pay, fey palmis ou lòt	0
	Tòl	4
	Beton	9
10. Eske fanmi an gen kochon?	Non	0
	Wi	5
		Total:

Translation provided by Anne Hasting and Julie Hubble of Fonkoze.

6. Scorecard use

As explained in Schreiner (2005), the main goal is not to maximize accuracy but rather to maximize the likelihood of programs' using scoring. When scoring projects fail, the culprit is usually not inaccuracy but rather the failure to convince users to accept scorecards and to use them properly (Schreiner, 2002). The roadblocks are less technical than human and organizational, less statistics than change management. "Accuracy" is easier—and matters less—than "practicality".

The simple, low-tech design here is meant to help users understand and trust the scorecard so that they will use it. While accuracy is important, it must be balanced against ease-of-use and "face validity". In particular, programs are more likely to collect data, compute scores, and pay attention to the results if, in their view, scoring avoids creating "extra" work and if the whole process generally seems to make sense to them.

This practical focus naturally leads to a one-page scorecard (Figure 2) that field workers can use to compute scores by hand in real time because it features:

- Few indicators
- Categorical indicators ("does the household own a pig", not "total value of assets")
- User-friendly weights (non-negative integers, no arithmetic beyond simple addition)

Among other things, this design permits rapid poverty appraisal, for example, determining in a day which village residents qualify for, say, microfinance membership or work-for-food programs.

The scorecard in Figure 2 could be photocopied to take to the field. It could also serve as a template for data-entry screens to record indicators, scores, poverty likelihoods, and changes in poverty likelihood over time.

When using the scorecard, field agents read each question, circle the response and the corresponding points, verify the response, write the points in the right-hand column, add up the points to get the score, and then execute program policy based on the score.

Field agents must be trained how to collect indicators. If they put garbage in, the scorecard will put garbage out. On-going audits of data quality are advisable.

Programs should record in a digital database at least the score, date, and participant identifier, if not also the values of the indicators. This will simplify computation of average poverty likelihoods and other analyses, both at a point in time and for changes through time (Matul and Kline, 2003).

7. Scores and poverty likelihoods

A *score* (sum of scorecard points) is not the same as a *poverty likelihood* (probability of being poor). But each score is associated with a poverty likelihood via a simple table (Figure 3, column “Poverty Likelihood for people with score in range (%)”).

For example, scores of 0–4 correspond to a poverty likelihood of 100.0 percent because, in the bootstrapped samples from the first hold-out sample from the 2001 ECVH (see below), everyone with scores of 0–4 were poor.

In the same way, scores of 45–49 correspond to a poverty likelihood of 14.3 percent, as this was the share of people (averaged over bootstraps) in the first hold-out sample who were poor.

In rough terms, the accuracy of scoring for targeting is the extent of concentration of the poor among low scores and of the non-poor among high scores.

In Figure 3, the column “% of people \leq score who are poor” shows the share of all Haitians with a given score or less who are poor. For example, 82.3 percent of those with scores of 20–24 or less are poor.

Likewise, the column “% of people $>$ score who are non-poor” shows the share of all Haitians with scores greater than a given range who are non-poor. For example, 67.1 percent of those with scores of more than 20–24 are poor.

Programs can use Figure 3 to set policy cut-offs for targeting program services. For example, suppose the program decides treat people scoring 24 or less as “poor” and people scoring 25 or more as “non-poor”. Then—assuming the program serves a population that mirrors that of Haiti as a whole—82.5 percent of those treated by the program as “poor” truly are poor (and 17.5 percent are non-poor), and 67.1 percent treated by the program as “non-poor” truly are non-poor (and 32.9 percent are poor).

Alternatively, the program could aim for a given overall poverty rate (say, 70 percent) and then choose a corresponding cut-off (here, 35–39). Usually, however, program participants will not mirror the population of Haiti as a whole, so the two right-hand columns of Figure 3 are not relevant.¹ A program aiming for a given overall poverty rate would set a cut-off and then monitor its overall poverty rate, adjusting the cut-off as required.

¹Even if participants do not mirror the country as a whole, Figure 3 can still be used with a net-benefit matrix to set policy cut-offs, as discussed later.

Figure 3: Scores and corresponding poverty likelihoods

Score	Poverty likelihood for people with score in range (%)	% of people <=score who are poor	% of people >score who are non-poor
0-4	100.0	100.0	47.1
5-9	90.9	93.2	49.8
10-14	83.0	88.0	53.4
15-19	82.8	85.9	59.6
20-24	74.5	82.5	67.1
25-29	60.9	77.9	74.2
30-34	50.7	73.8	80.8
35-39	40.1	69.9	87.0
40-44	18.9	65.3	89.1
45-49	14.3	62.4	90.1
50-54	11.7	59.6	90.9
55-59	14.0	57.8	93.3
60-64	4.6	56.3	92.1
65-69	0.0	55.2	87.0
70-74	23.0	54.7	100.0
75-79	0.0	54.1	100.0
80-84	0.0	54.1	100.0
85-89	0.0	54.0	0.0
90-94	0.0	54.0	0.0
95-100	0.0	54.0	N/A
Total:	54.0		

Surveyed cases weighted to represent all Haiti.

Source: Calculations by Microfinance Risk Management, L.L.C., based on 2001 ECVH.

8. Correspondences between scores and poverty likelihoods

The poverty likelihoods in Figure 3 are derived from the 2001 ECVH using a bootstrapped hold-out sample.

At the start of the study, two *hold-out samples* of 1,434 households each (20 percent of the 2001 ECVH in each hold-out sample) was selected at random, and all steps in scorecard construction were done on the remaining 4,300 households (without “peeking” at the hold-out samples). The hold-out samples were then used to determine poverty likelihoods and to measure precision.

Except Fuller (2006), Schreiner (2006e), and Setel *et al.* (2003), all poverty scorecards to date have been built and tested on the same set of households. This overstates accuracy because all scorecards are inevitably overfit to some extent.² *Overfit* means that the choice of indicators, the form of indicators, and the weights represent not only universal, permanent patterns present among all households but also random and/or transitory patterns found only among the households used to build the scorecard.

To get measures of precision uncontaminated by overfitting requires testing on households not used to build the scorecard. This also mimics how the scorecard is actually used in practice.

Poverty likelihoods are derived from the first hold-out sample as follows:

- Score all households in the hold-out sample
- For a given score, define the *poverty likelihood* as the share of people with that score who are poor

For example, suppose that 82.8 percent of people in the hold-out sample with scores of 15–19 are poor. The poverty likelihood for scores of 15–19 is then 82.8 percent.

Of course, drawing a different hold-out sample would lead to a different poverty likelihood. While estimates vary from sample to sample, a precise estimator is one that usually is close to the true value being estimated.

² Bigman *et al.* (2000) show this for a poverty scorecard in Burkina Faso.

Bootstrapping is a simple way to measure precision (Efron and Tibshirani, 1993). In this paper, it is also used to produce estimates of poverty likelihoods that are free of overfitting. The algorithm is:

- From the hold-out sample, draw a new sample of 1,434 households *with replacement*³
- For each given range of scores, compute the poverty likelihood as described above
- Repeat the previous two steps many times (here, 10,000)
- For a given score range, define the poverty likelihood (first column of Figure 3) as the average of the 10,000 poverty likelihoods in the bootstrapped samples

Precision is expressed as confidence intervals (Figure 4), a standard statistical technique that is well-understood by the general public.⁴

For example, the average across all bootstrap samples of the share of people with scores of 15–19 who were poor was 82.8 percent. In 90 percent of the 10,000 bootstrapped samples, the share was between 78.3–89.6 percent, an interval of about ± 6.5 percentage points.

The 90-percent confidence intervals are consistently about 6.5 percentage points for all score ranges in which many people fall. They widen for very low or very high scores because few people fall in these ranges.

Narrower confidence intervals mean greater precision. Among other things, the width of confidence intervals depends on the size of the hold-out sample, the number of people with a given score, the accuracy of the scorecard, and the extent of overfitting.

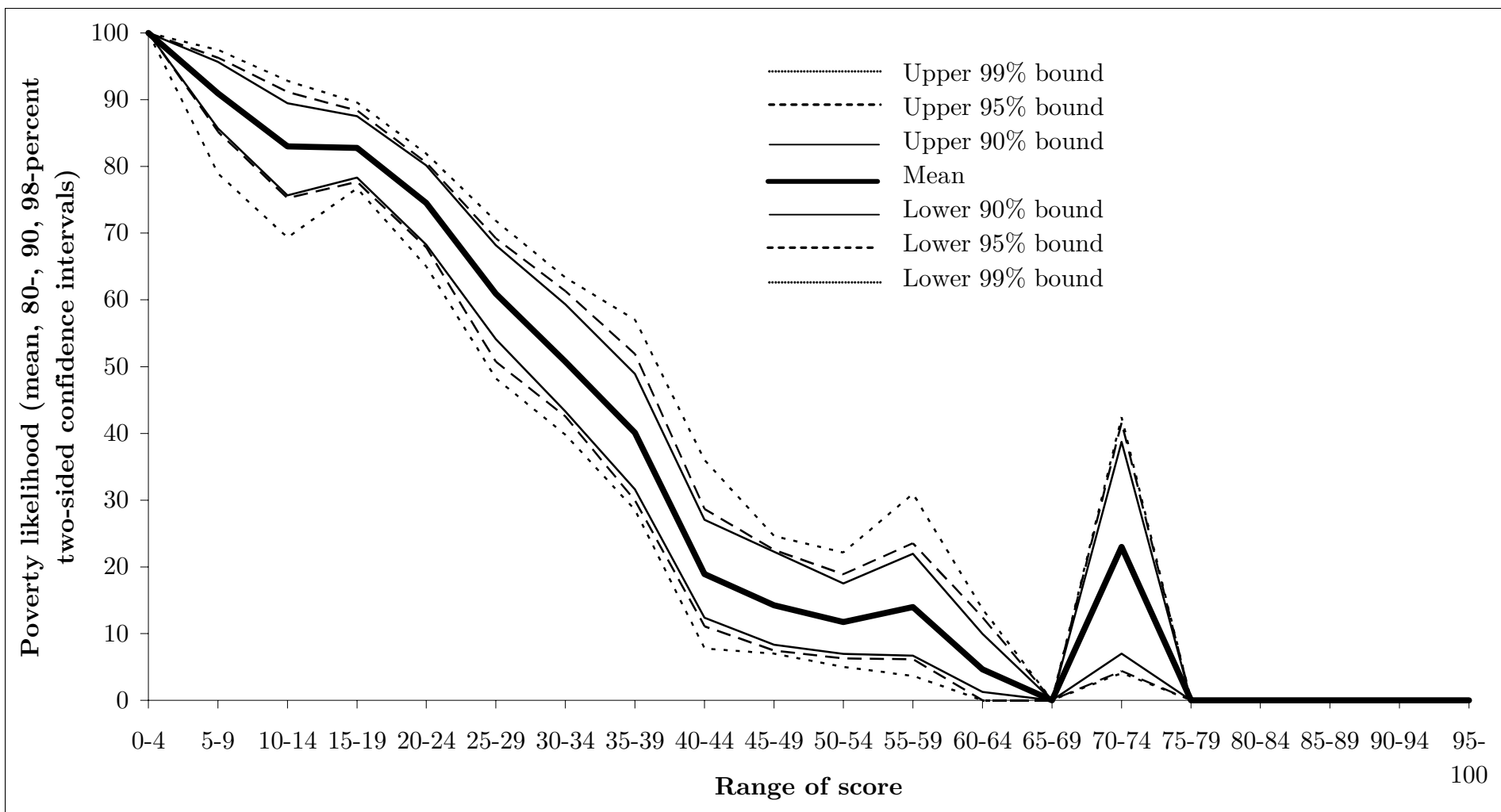
There is no absolute benchmark for what is “precise enough”. If other scorecards measured precision with confidence intervals, they could be compared.

Note that this scorecard produces objective (data-based) estimates of poverty likelihood. This holds even though some qualitative judgment is used—along with “MAXC”—to select indicators. In fact, objective scorecards of proven accuracy are often constructed based *only* on qualitative judgment (Caire, 2004; Schreiner *et al.*, 2004; Lovie and Lovie, 1986; Dawes, 1979; Wainer, 1976). What makes for objectivity is not how scorecards are constructed but rather how scores are linked with poverty likelihoods.

³ Because samples are drawn with replacement, the probability that any particular household will appear at least once in a given bootstrap sample is about 63.2 percent.

⁴ This was first done for poverty scorecards (a.k.a. “proxy means tests”) in Schreiner (2006e). For examples of its application in small-area poverty mapping, see Elbers, Lanjouw, and Lanjouw (2003), and Hentschel *et al.* (2000).

Figure 4: Confidence intervals for poverty likelihoods



9. Estimates of poverty rates

The overall poverty rate for all participants—the number that USAID microenterprise grantees must report—is the average of the poverty likelihoods of all participants.

For example, suppose a Haitian development program had three participants on Jan. 1, 2006 with scores of 20, 30, and 40, corresponding to poverty likelihoods of 74.5, 50.7, and 18.9 percent. The overall poverty rate is the participants' average poverty likelihood, that is, $(74.5 + 50.7 + 18.9) \div 3 = 48.0$ percent.

The precision of the estimated poverty rate was measured by drawing 10,000 new bootstrap samples from the second hold-out sample. The distribution of the 10,000 differences between the scorecard's estimate of the overall poverty rate (average poverty likelihood) and the true poverty rate is—in line with theory—Normal ($\mu = -0.0076$, $\sigma = 0.0115$).

The scorecard's average estimate is about 0.8 percentage points too low. If all the assumptions of Logit regression hold, the estimator should be unbiased, but at least two assumptions do not hold:

- The weights are not exactly optimal due to rounding in the transformation from the original Logit weights into non-negative integers whose sum is between 0 and 100
- Failure of the scorecard to include a complete set of all relevant indicators⁵

This bias has a simple remedy; estimate the overall poverty rate as the average poverty likelihood, plus 0.8 percentage points.

How precise are the scorecard estimates? The bootstrap indicates that there is 90-percent confidence that the true poverty rate is within ± 1.9 percentage points of the estimate, 95-percent confidence for ± 2.3 percentage points, and 99-percent confidence for ± 3.0 percentage points. For program purposes, this level of precision is probably more than adequate.

⁵ Of course, this omitted-variable bias is ubiquitous in any scoring exercise. In contrast, the estimators in IRIS (2005a) and Zeller, Alcaraz V., and Johannsen (2004) are biased even if all their modeling assumptions hold.

10. Progress out of poverty through time

For a given group, progress out of poverty over time is estimated as the change in average poverty likelihood.

Continuing the example from the previous section, suppose that on Jan. 1, 2007, the same three people (some of whom may no longer be participants) have scores of 25, 35, and 60 (poverty likelihoods of 60.9, 40.1, and 4.6 percent). Their average poverty likelihood is now 35.2 percent, an improvement of $48.0 - 35.2 = 12.8$ percentage points.

In a large portfolio, this means 12.8 of every 100 participants exited poverty. Given that 54.0 percent of participants were poor in the first place, about one in four ($13.5 \div 54.0 = 25.0$ percent) of poor participants left poverty.

Of course, this does not mean that participation in the development program *caused* the progress; the scorecard just measures what happened, regardless of cause.

11. Accuracy in targeting

While accuracy is not the only (nor main) goal, it is important. The individual poverty likelihoods and the overall poverty rate are accurate by their construction from bootstrapping. When using scoring for targeting, greater accuracy means that the poor are more concentrated in low scores and the non-poor in high scores.

At the extreme, a perfect scorecard would assign all the lowest scores to poor people and all the highest scores to non-poor people, for example, if everyone with a score of 49 or less were poor and everyone with a score of 50 or more were non-poor. In reality, no scorecard is perfect. Some non-poor people have low scores, and vice versa.

A person has one of two *poverty statuses*:

- Poor: Expenditure at or below the poverty line
- Non-poor: Expenditure above the poverty line

Poverty status is a fact. If there is data on income (as in the ECVH), then poverty status is known.

A person can also be classified into one of two *poverty segments*:

- Poor: Score at or below a poor/non-poor cut-off
- Non-poor: Score above a poor/non-poor cut-off

Poverty segment is program-determined. For example, a program might set a cut-off of 20–24. *For program purposes*, people with scores at or below 20–24 are treated as if they were poor, and the rest are treated as non-poor.

Because no scorecard is perfect, *poverty status* (expenditure vis-à-vis a poverty line) sometimes differs from *poverty segment* (score vis-à-vis a program's cut-off). That is, some people whose status is truly poor are classified in the non-poor segment, and vice versa. Targeting is accurate to the extent that poverty segment matches poverty status.

Programs use poverty segment as a proxy for (unknown) poverty status. Indeed, the purpose of scoring is to infer poverty status without having to measure expenditure. For people in the 2001 ECVH, both poverty status and poverty segment are known, so their coincidence (targeting accuracy) can be measured.

Suppose that a program defines a poor/non-poor cut-off of 20–24. According to the column “% of people \leq score who are poor” in Figure 3, 82.5 percent of Haitians with scores at or below 20–24 were truly poor. (The other 17.5 percent were non-poor.)

At the same time, the column “% of people >score who are non-poor” shows that 67.1 percent of those with scores of more than the 20–24 cut-off were truly non-poor (and thus 32.9 percent were poor). In sum, a poor/non-poor cut-off of 20–24 correctly classifies (that is, poverty status matches poverty segment) 82.5 percent of the people classified as poor and 67.1 percent of the people classified as non-poor.

How does classification accuracy depend on the cut-off? Using 30–34 (rather than 20–24) correctly classifies 50.7 percent of those classified as poor and 80.8 percent of those classified as non-poor. This illustrates a general point; better targeting for the poor comes at the cost of worse targeting for the non-poor (and vice versa).

12. Setting the poor/non-poor cut-off

To choose a cut-off, programs need a way to trade off accuracy for the poor versus accuracy for the non-poor. The standard way uses a *classification matrix* and a *net-benefit matrix*.

Classification matrix

Given a poor/non-poor cut-off, there are four types of classification results:

- A. Truly poor correctly classified in poor segment (score at or below the cut-off)
- B. Truly poor incorrectly classified in non-poor segment (score above cut-off)
- C. Truly non-poor incorrectly classified in poor segment (score at or below cut-off)
- D. Truly non-poor correctly classified in non-poor segment (score above cut-off)

These four results can be thought of as a classification matrix:

Figure 5: General classification matrix

		<u>Poverty segment</u>	
		<u>Poor</u>	<u>Non-poor</u>
<u>True poverty status</u>	<u>Poor</u>	A. Truly poor correctly classified in poor segment	B. Truly poor incorrectly classified in non-poor segment
	<u>Non-poor</u>	C. Truly non-poor incorrectly classified in poor segment	D. Truly non-poor correctly classified in non-poor segment

Accuracy improves as greater shares fall in quadrants A and D and fewer in B and C.

Figure 6 is the share of Haitians in each classification for all cut-offs. For 20–24:

- A. 35.1% are correctly classified as poor (truly poor in poor segment)
- B. 18.9% are incorrectly classified as non-poor (truly poor in non-poor segment)
- C. 7.4% are incorrectly classified as poor (truly non-poor in poor segment)
- D. 38.6% are correctly classified as non-poor (truly non-poor in non-poor segment)

Figure 6: Share of people by classification

Score	A.	B.	C.	D.
	Truly poor correctly classified in poor segment	Truly poor incorrectly classified in non-poor segment	Truly non-poor incorrectly classified in poor segment	Truly non-poor correctly classified in non-poor segment
0-4	2.2	51.8	0.0	46.0
5-9	8.1	45.9	0.6	45.4
10-14	15.6	38.3	2.1	43.9
15-19	25.6	28.4	4.2	41.8
20-24	35.1	18.9	7.4	38.6
25-29	42.2	11.8	12.0	34.1
30-34	47.0	7.0	16.7	29.4
35-39	50.3	3.6	21.7	24.4
40-44	51.7	2.3	27.5	18.5
45-49	52.4	1.6	31.6	14.4
50-54	52.9	1.0	35.9	10.1
55-59	53.5	0.5	39.1	7.0
60-64	53.6	0.4	41.5	4.5
65-69	53.6	0.4	43.5	2.6
70-74	54.0	0.0	44.8	1.3
75-79	54.0	0.0	45.9	0.2
80-84	54.0	0.0	45.9	0.2
85-89	54.0	0.0	46.0	0.0
90-94	54.0	0.0	46.0	0.0
95-100	54.0	0.0	46.0	0.0

Source: Calculations by Microfinance Risk Management, L.L.C., based on 2001 ECVH.

Figures normalized to sum to 100.

If the cut-off rises to 25–29, more poor (but less non-poor) are correctly classified:

- A. 42.2% are correctly classified as poor (truly poor in poor segment)
- B. 11.8% are incorrectly classified as non-poor (truly poor in non-poor segment)
- C. 12.0% are incorrectly classified as poor (truly non-poor in poor segment)
- D. 34.1% are correctly classified as non-poor (truly non-poor in non-poor segment)

Whether a cut-off of 20–24 is preferred to a cut-off of 25–29 depends on net benefit.

Net-benefit matrix

Each of the four types of classification results is associated with a net benefit:

- α . Benefit of truly poor correctly classified in poor segment
- β . Cost (negative net benefit) of truly poor incorrectly classified in non-poor segment
- γ . Cost (negative net benefit) of truly non-poor incorrectly classified in poor segment
- δ . Benefit of truly non-poor correctly classified in non-poor segment

Figure 7: General net-benefit matrix

		<u>Poverty segment</u>	
		<u>Poor</u>	<u>Non-poor</u>
<u>True poverty status</u>	<u>Poor</u>	α	β
	<u>Non-poor</u>	γ	δ

Given a net-benefit matrix and a classification matrix, total net benefit is:

$$\text{Total net benefit} = \alpha \cdot A + \beta \cdot B + \gamma \cdot C + \delta \cdot D.$$

To choose the optimal cut-off, a program would:

- Define a net-benefit matrix based on the program’s values and mission
- Compute total net benefits for each cut-off using Figure 6 and the net-benefit matrix
- Select the cut-off with the highest total net benefit

Most pro-poor development programs care about correctly classifying both the poor and non-poor, even if the poor matter more. Thus, most programs will have non-zero values in at least three of the four quadrants of the net-benefit matrix. The use of a net-benefit matrix allows programs to be explicit and intentional about how they value all the trade-offs inherent when setting cut-offs. This is why the use of classification matrices and net-benefit matrices is standard in scoring (SAS, 2004; SPSS, 2003; Adams and Hand, 2000; Salford Systems, 2000; Hoadley and Oliver, 1998; Greene, 1993).

“Total Accuracy”

For example, suppose a program selects the net-benefit matrix that corresponds to the “Total Accuracy” criterion (IRIS, 2005b).

Figure 8: “Total Accuracy” net-benefit matrix

		<u>Poverty segment</u>	
		<u>Poor</u>	<u>Non-poor</u>
<u>True poverty status</u>	<u>Poor</u>	1	0
	<u>Non-poor</u>	0	1

With “Total Accuracy”, total net benefit is the number of people correctly classified:

$$\begin{aligned} \text{Total net benefit} &= 1 \cdot A + 0 \cdot B + 0 \cdot C + 1 \cdot D, \\ &= A + D. \end{aligned}$$

Grootaert and Braithwaite (1998) and Zeller, Alcaraz, and Johannsen (2004) use “Total Accuracy” as a measure of scorecard accuracy.

Figure 9 shows “Total Accuracy” for all cut-offs. Total net benefit is highest (76.3) with a cut-off of 30–34; here, poverty segment matches poverty status for about three in four Haitians.

A weakness of “Total Accuracy” is that it weighs correct classifications of the poor and non-poor equally (IRIS, 2005b). If most people are non-poor and/or if a scorecard is more accurate for the non-poor, then “Total Accuracy” might be high even if few poor people are correctly classified. Programs targeting the poor, however, probably value correct classification more for the poor than the non-poor.

Figure 9: Net benefits for common net-benefit matrices

Score	<u>Total Accuracy</u> (A + B)		<u>Poverty Accuracy</u> $100*A / (A+B)$		<u>Non-poverty Accuracy</u> $100*D / (C+D)$		<u>Undercoverage</u> $100*B / (A+B)$		<u>Leakage</u> $100*C / (A+C)$	
	1	0	1	0	0	0	0	-1	0	0
	0	1	0	0	0	1	0	0	-1	0
0-4	48.2		4.1		100.0		95.9		0.0	
5-9	53.5		15.0		98.7		85.0		6.8	
10-14	59.5		28.9		95.4		71.1		12.0	
15-19	67.4		47.5		90.9		52.5		14.1	
20-24	73.7		65.0		83.9		35.0		17.5	
25-29	76.2		78.1		74.0		21.9		22.1	
30-34	76.3		87.1		63.8		12.9		26.2	
35-39	74.7		93.3		52.9		6.7		30.1	
40-44	70.2		95.8		40.3		4.2		34.7	
45-49	66.8		97.0		31.4		3.0		37.6	
50-54	63.1		98.1		22.0		1.9		40.4	
55-59	60.4		99.1		15.1		0.9		42.2	
60-64	58.1		99.3		9.7		0.7		43.7	
65-69	56.1		99.3		5.6		0.7		44.8	
70-74	55.2		100.0		2.8		0.0		45.3	
75-79	54.1		100.0		0.4		0.0		45.9	
80-84	54.1		100.0		0.4		0.0		45.9	
85-89	54.0		100.0		0.0		0.0		46.0	
90-94	54.0		100.0		0.0		0.0		46.0	
95-100	54.0		100.0		0.0		0.0		46.0	

All figures in percentage units.

A simple, transparent way to reflect this is to increase the relative net benefit of correctly classifying the poor. For example, if a program values correctly classifying the poor twice as much as correctly classifying the non-poor, then α should be set twice as high as δ in the net-benefit matrix. Then the new optimal cut-off is 35–39, the point where $2 \cdot A + D$ is highest.

“Poverty Accuracy”

A criterion that emphasizes the importance of correctly classifying the poor is “Poverty Accuracy” (IRIS, 2005b).

Figure 10: “Poverty Accuracy” net-benefit matrix

		<u>Poverty segment</u>	
		<u>Poor</u>	<u>Non-poor</u>
<u>True poverty status</u>	<u>Poor</u>	1	0
	<u>Non-poor</u>	0	0

“Poverty Accuracy” only counts correct classifications of the poor:

$$\begin{aligned} \text{Total net benefit} &= 1 \cdot A + 0 \cdot B + 0 \cdot C + 0 \cdot D, \\ &= A. \end{aligned}$$

The weakness is that correct classification of the poor is rarely the sole criteria. In fact, Figure 9 shows that “Poverty Accuracy” is always maximized with a cut-off of 95–100. While classifying everyone as poor does ensure that all poor people qualify for program services and thus minimizes *undercoverage* of the poor (second-to-last column of Figure 9), it also maximizes *leakage* (the last column), as all non-poor people are also classified as poor.

In short, maximizing “Poverty Accuracy” means universal programs (no targeting). In some contexts, this is appropriate; the point here is to make explicit the implications of “Poverty Accuracy” as a criterion for choosing a poor/non-poor cut-off.

“Non-poverty Accuracy”

“Non-poverty Accuracy” counts only correct classifications of the non-poor (total net benefit is D). Of course, this is maximized by setting a cut-off of 0–4 so that everyone is classified as non-poor. This is not useful, as it means no one is targeted for program services (leakage is minimized, but undercoverage is maximized).

“BPAC”

IRIS (2005b) proposes a new measure of scorecard accuracy called the “Balanced Poverty Accuracy Criterion” (BPAC). It attempts to balance two goals:

- Maximize the accuracy of the estimated overall poverty rate
- Maximize “Poverty Accuracy”

For the first goal, the estimated poverty rate is most accurate when undercoverage B equals leakage C. For the second goal, “Poverty Accuracy” is best when A is maximized. If $B > C$, then the implicit net-benefit matrix for “BPAC” is:

Figure 11: “BPAC” net-benefit matrix

		<u>Poverty segment</u>	
		<u>Poor</u>	<u>Non-poor</u>
<u>True poverty status</u>	<u>Poor</u>	1	1
	<u>Non-poor</u>	-1	0

If $B > C$, then “BPAC” maximizes A while making B as close to C as possible:

$$\begin{aligned} \text{Total net benefit} &= 1 \cdot A + 1 \cdot B + (-1) \cdot C + 0 \cdot D, \\ &= A + (B - C). \end{aligned}$$

If $C > B$, then total net benefit under BPAC is $A + (C - B)$.

Unfortunately, BPAC is not meaningful for scorecards that estimate poverty likelihoods rather than expenditure (Schreiner, 2005). Instead, this paper takes the standard, well-understood approach to measuring accuracy and precision via the statistical concepts of “bias” (in repeated samples, how close on average the estimate is to the true value) and “confidence intervals” (in repeated samples, how often the estimate falls within a given distance of the true value). Rather than “BPAC is x ”, the measures are “There is x -percent confidence that the estimated overall poverty rate is within $\pm y$ percentage points of the true value”.

Summary of accuracy discussion

A poverty scorecard is used for:

- Estimating overall poverty rates
- Estimating individual poverty likelihoods
- Classifying people for targeting purposes

Estimates are accurate to the extent that they match the true value being estimated.

For overall poverty rates, estimates from the Haiti scorecard have a 90-percent chance of being within 1.5 percentage points of the true poverty rate. For individual poverty likelihoods, estimates have a 90-percent chance of being within 1.9 percentage points of the true poverty rate.

For targeting with individual classifications, accuracy varies by scorecard and by the poor/non-poor policy cut-offs defined by the program. The most appropriate measure is total net benefit based on a program-specific net-benefit matrix and the classification results in Figure 6. Total net benefit is not an absolute benchmark, but a given program can use it to choose between two different scorecards.

A general, non-program-specific measure of targeting accuracy is “ c ”, the share of all pairs of poor and non-poor households in which the poor household has a lower score. For this scorecard, “ c ” is 77.1 percent.

13. Three classes: very poor, poor, and non-poor

Congress requires USAID microfinance grantees to report the share of participants who are poor versus non-poor. Grameen Foundation U.S.A., however, looks not only at the poor and non-poor but also at the *very poor*, that is, the poorest half of the poor.

For GFUSA, a person has one of three poverty statuses:

- Very poor: Poorest half of those with expenditure at or below the poverty line
- Poor: Least-poor half of those with expenditure at or below poverty
- Non-poor: Expenditure above poverty

There are also three poverty segments:

- Very poor: Score at or below the very poor/poor cut-off
- Poor: Score above the very poor/poor cut-off and at or below the poor/non-poor cut-off
- Non-poor: Score above the poor/non-poor cut-off

There are two cut-offs (very poor/poor and poor/non-poor) and 9 classification results:

A. Truly very poor	correctly classified	as very poor
B. Truly very poor	incorrectly classified	as poor
C. Truly very poor	incorrectly classified	as non-poor
D. Truly poor	incorrectly classified	as very poor
E. Truly poor	correctly classified	as poor
F. Truly poor	incorrectly classified	as non-poor
G. Truly non-poor	incorrectly classified	as very poor
H. Truly non-poor	incorrectly classified	as poor
I. Truly non-poor	correctly classified	as non-poor

Rather than use choosing policy cut-offs using the explicit, rational paradigm of classification matrices and net-benefit matrices, GFUSA has opted instead to leave the choice process and criteria implicit. Thus, a discussion of this paradigm when there are three poverty statuses is omitted. It may be found in Schreiner (2006c and 2006d).

As in those previous documents, it is hereby noted that it is inappropriate to have anyone (such as a consultant) other than the program set their policy cut-offs (of course, experts can assist programs to work through the appropriate process). In particular, it is in general inappropriate to simply set cut-offs so that individuals are assigned to the class for which they have the highest poverty likelihood.

Figure 12: Classification results, 3 classes (very poor, poor, non-poor)

Score	% people with score in range			% people \leq score who are:		
	Very Poor	Poor	Non-poor	Very Poor	Poor	Non-poor
0-4	88.9	11.1	0.0	88.9	11.1	0.0
5-9	61.1	29.9	9.1	68.2	25.1	6.8
10-14	42.9	40.1	17.0	55.2	32.8	12.0
15-19	44.2	38.6	17.2	50.8	35.1	14.1
20-24	45.9	28.7	25.5	49.3	33.2	17.5
25-29	26.5	34.4	39.1	44.4	33.5	22.1
30-34	21.0	29.8	49.3	40.9	32.9	26.2
35-39	11.8	28.3	59.9	37.5	32.4	30.1
40-44	11.7	7.3	81.1	35.2	30.1	34.7
45-49	1.0	13.2	85.7	33.2	29.1	37.6
50-54	1.8	9.9	88.3	31.5	28.1	40.4
55-59	0.8	13.2	86.0	30.3	27.5	42.2
60-64	0.0	4.6	95.4	29.5	26.9	43.7
65-69	0.0	0.0	100.0	28.9	26.3	44.8
70-74	1.0	22.0	77.0	28.4	26.3	45.3
75-79	0.0	0.0	100.0	28.1	26.0	45.9
80-84	0.0	0.0	100.0	28.1	26.0	45.9
85-89	0.0	0.0	100.0	28.0	25.9	46.0
90-94	0.0	0.0	100.0	28.0	25.9	46.0
95-100	0.0	0.0	100.0	28.0	25.9	46.0
Total:	28.0	25.9	46.0			

In Figure 12, the three left-side columns under “% people with score in range” link scores with poverty likelihoods. Among people with scores of 20–24, for example, 45.9 percent are very poor, 28.7 percent are poor, and 25.5 percent are non-poor. The scorecard has good accuracy, as the very poor are concentrated in the lowest scores, the poor in “middle” scores, and the non-poor in the highest scores.

The three right-side columns in Figure 12 under “% people \leq score who are:” show the poverty composition of a portfolio if a program accepts only participants at or below a given cut-off (and if applicants mirror the population for Haiti as a whole). A cut-off of 20–24, for example, leads to a portfolio in which 49.3 percent of participants are very poor, 33.2 percent are poor, and 17.5 percent are non-poor.

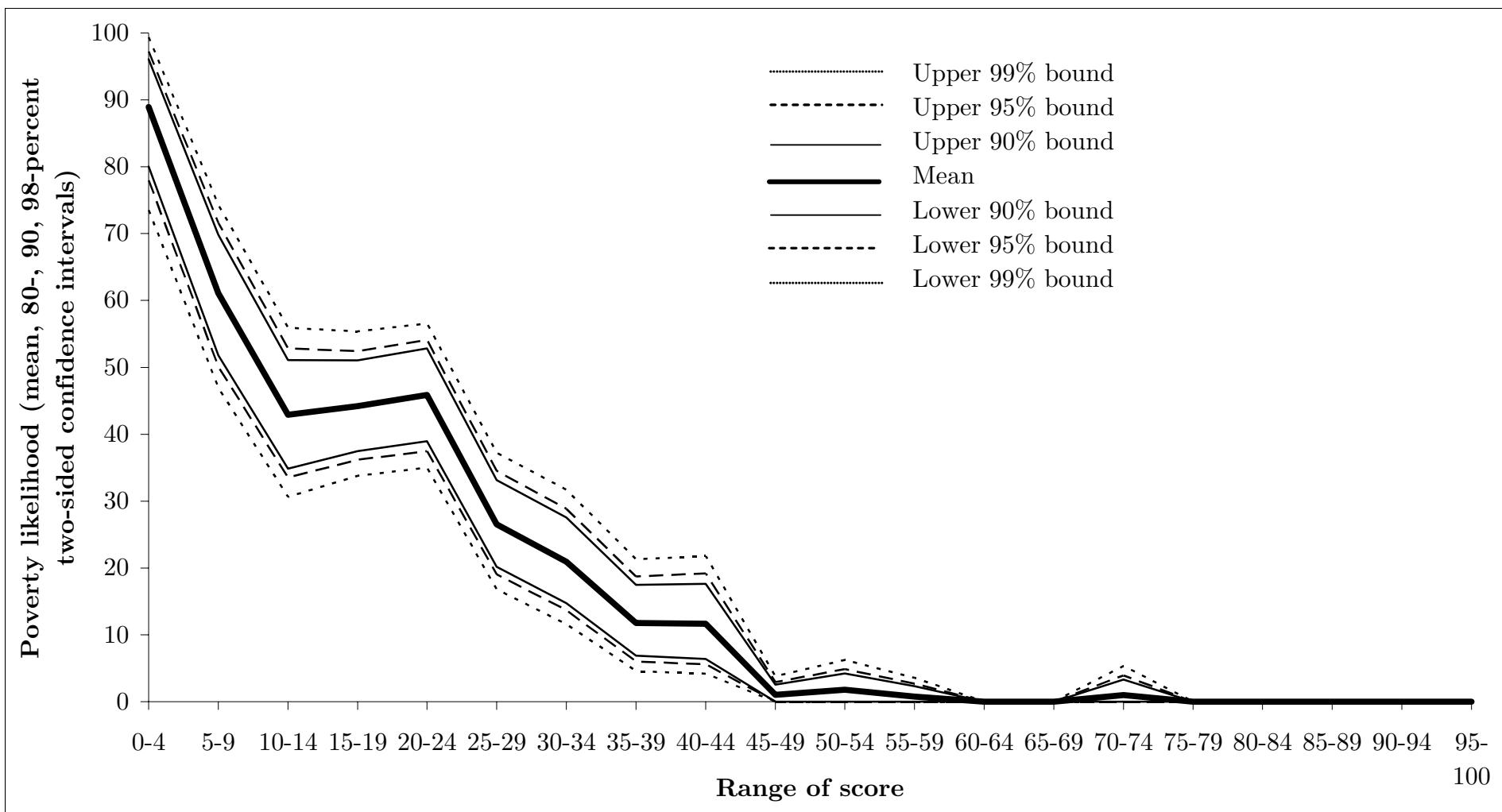
Programs can thus use Figure 12 to set policy cut-offs. Of course, their applicant pool probably does not mirror the population of Haiti as a whole, but programs can, as a first cut, assume that it does and then monitor their actual overall poverty rates, adjusting the cut-offs up or down over time as needed.

As in the two-class case, poverty likelihoods are derived from bootstrapping the second hold-out sample. Figure 13 depicts 90-, 95-, and 99-percent confidence intervals for the estimated poverty likelihoods. In the range of scores in which there are large numbers of very poor people, the 90-percent confidence interval is about ± 7.0 percentage points.

The estimated “very poor” poverty rate is the average “very poor” poverty likelihood among a group of individuals. On average in 10,000 bootstrap samples, the distribution of the difference between the estimated “very poor” poverty rate and the true “very poor” poverty rate was Normal ($\mu = 0.0040$, $\sigma = 0.0110$).

Thus, the best estimate of the overall “very poor” poverty rate is the average “very poor” poverty likelihood, minus 0.4 percentage points. The 90-percent confidence interval is ± 1.8 percentage points, the 95-percent confidence interval is ± 2.2 percentage points, and the 99-percent confidence interval is ± 2.8 percentage points.

Figure 13: Confidence intervals for poverty likelihoods associated with being very poor



14. Summary

- Haiti is the poorest country in the Western Hemisphere, and the most unequal. A simple, easy-to-use, inexpensive tool for identifying the poor could improve the targeting of pro-poor development programs and help speed progress out of poverty
- The scorecard here estimates the likelihood that a person has expenditure of less than \$1/day. It estimates accurately:
 - The likelihood that an individual is poor (within ± 6.5 percentage points with 90-percent confidence)
 - The overall poverty rate (within ± 1.9 percentage points with 90-percent confidence)
- Accuracy is objectively proven, as scores are related to poverty likelihoods via the 2001 ECVH. Precision is measured via bootstrapping on two hold-out samples
- Field workers can compute scores on paper in real time
- The scorecard can be used by any program seeking a quick, easy, inexpensive, and accurate way to identify the poor
- The scorecard can be used to accurately classify people as very poor/poor/non-poor
- Overall, pro-poor development programs in Haiti can use the simple poverty scorecard presented in this paper to:
 - Target services to the poor
 - Track participants' progress out of poverty through time
 - Report on the share of participants are poor

References

- Adams, N.M.; and D.J. Hand. (2000) "Improving the Practice of Classifier Performance Assessment", *Neural Computation*, Vol. 12, pp. 305–311.
- Baulch, Bob. (2003) "Poverty Monitoring and Targeting Using ROC Curves: Examples from Vietnam", IDS Working Paper No. 161, <http://www.ids.ac.uk/ids/bookshop/wp/wp161.pdf>.
- Bigman, David; Dercon, Stefan; Guillaume, Dominique; and Michel Lambotte. (2000) "Community Targeting for Poverty Reduction in Burkina Faso", *World Bank Economic Review*, Vol. 14, No. 1, pp. 167–193.
- Caire, Dean. (2004) "Building Credit Scorecards for Small Business Lending in Developing Markets", Bannock Consulting, http://www.microfinance.com/English/Papers/Scoring_SMEs_Hybrid.pdf.
- Dawes, Robyn M. (1979) "The Robust Beauty of Improper Linear Models in Decision Making", *American Psychologist*, Vol. 34, No. 7, pp. 571–582.
- Efron, Bradley; and Robert J. Tibshirani. (1993) *An Introduction to the Bootstrap*, New York, NY: Chapman and Hall, ISBN 0–412–04231–2.
- Elbers, Chris; Lanjouw, J.O.; and Peter Lanjouw (2003) "Micro-Level Estimation of Poverty and Inequality," *Econometrica*, Vol. 71, No. 1, pp. 355–64.
- Fuller, Rob. (2006) "Poverty Indicators for Fonkoze Clients: Benchmarking the *Kat Evalyasyon*", Research and Social Impact Department, rfuller@fonkoze.org.
- Goodman, L.A. and Kruskal, W.H. (1979) *Measures of Association for Cross Classification*, New York, NY: Springer-Verlag, ISBN 0–38–790443–3.
- Greene, William H. (1993) *Econometric Analysis: Second Edition*, New York, NY: MacMillan, ISBN 0–02–346391–0.
- Grootaert, Christiaan; and Jeanine Braithwaite. (1998) "Poverty Correlates and Indicator-Based Targeting in Eastern Europe and the Former Soviet Union", World Bank Policy Research Working Paper No. 1942, Washington, D.C., <http://www.worldbank.org/html/dec/Publications/Workpapers/WPS1900series/wps1942/wps1942.pdf>.

- Hentschel, Jesko; Olsen Lanjouw, Jean; Lanjouw, Peter; and Javier Poggi. (2000) “Combining Census and Survey Data to Trace the Spatial Dimensions of Poverty: A Case Study of Ecuador”, *World Bank Economic Review*, Vol. 14, No. 1, pp. 147–165.
- Hoadley, Bruce; and Robert M. Oliver. (1998) “Business measures of scorecard benefit”, *IMA Journal of Mathematics Applied in Business and Industry*, Vol. 9, pp. 55–64.
- IRIS Center. (2005a) “Accuracy Results for 12 Poverty Assessment Tool Countries”, <http://www.povertytools.org/documents/Accuracy%20Results%20for%2012%20Countries.pdf>.
- IRIS Center. (2005b) “Notes on Assessment and Improvement of Tool Accuracy”, <http://www.povertytools.org/documents/Assessing%20and%20Improving%20Accuracy.pdf>.
- Lovie, A.D.; and P. Lovie. (1986) “The Flat Maximum Effect and Linear Scoring Models for Prediction”, *Journal of Forecasting*, Vol. 5, pp. 159–168.
- Matul, Michal; and Sean Kline. (2003) “Scoring Change: Prizma’s Approach to Assessing Poverty”, MFC Spotlight Note No. 4, Warsaw, Poland: Microfinance Centre for Central and Eastern Europe and the New Independent States, http://www.mfc.org.pl/doc/Research/ImpAct/SN/MFC_SN04_eng.pdf.
- Pedersen, Jon; and Kathryn Lockwood. (2001) “Determination of a Poverty Line for Haiti”, FAFO Institute of Applied International Studies, <http://www.fafo.no/ais/other/haiti/poverty/PovertyLineForHaiti.pdf>.
- Salford Systems. (2000) *CART for Windows User’s Guide*, San Diego, CA.
- SAS. (2004) *SAS OnlineDoc 9.1.2*, Cary, NC.
- Schreiner, Mark. (2006a) “Indices de Pobreza para México”, report to Grameen Foundation U.S.A., http://www.microfinance.com/Castellano/Documentos/Scoring_Poverty_Mexico.pdf.
- Schreiner, Mark. (2006b) “A Simple Poverty Scorecard for the Philippines”, report to Grameen Foundation U.S.A., http://www.microfinance.com/English/Papers/Scoring_Poverty_Pakistan.pdf.

- Schreiner, Mark. (2006c) “A Simple Poverty Scorecard for India”, report to Grameen Foundation U.S.A.,
http://www.microfinance.com/English/Papers/Scoring_Poverty_India.pdf.
- Schreiner, Mark. (2006d) “A Simple Poverty Scorecard for Pakistan”, report to Grameen Foundation U.S.A.,
http://www.microfinance.com/English/Papers/Scoring_Poverty_Pakistan.pdf.
- Schreiner, Mark. (2006e) “A Simple Poverty Scorecard for Pakistan”, report to Grameen Foundation U.S.A.,
http://www.microfinance.com/English/Papers/Scoring_Poverty_Bangladesh.pdf.
- Schreiner, Mark. (2005) “IRIS questions on poverty scorecards”, memo for Grameen Foundation U.S.A., mark@microfinance.com.
- Schreiner, Mark. (2002) *Scoring: The Next Breakthrough in Microfinance?* Occasional Paper No. 7, Consultative Group to Assist the Poorest, Washington, D.C.,
http://www.cgap.org/docs/OccasionalPaper_07.pdf,
- Schreiner, Mark; Matul, Michal; Pawlak, Ewa; and Sean Kline. (2004) “Poverty Scorecards: Lessons from a Microlender in Bosnia-Herzegovina”, Microfinance Risk Management,
http://www.microfinance.com/English/Papers/Scoring_Poverty_in_BiH_Short.pdf.
- Setel, Philip; Abeyasekera, Savitri; Ward, Patrick; Hemed, Yusuf; Whiting, David; Mswia, Robert; Antoninis, Manos; and Henry Kitange. (2003) “Development, Validation, and Performance of a Rapid Consumption Expenditure Proxy for Measuring Income Poverty in Tanzania: Experience from AMMP Demographic Surveillance Sites”,
http://www.rdg.ac.uk/ssc/publications/Inc_PovProx.pdf.
- Sletten, Pål; and Willy Egset. (2004) “Poverty in Haiti”, FAFO Paper No. 2004:31,
<http://www.fafo.no/PUB/rapp/755/755.pdf>.
- SPSS, Inc. (2003) *Clementine 8.0 User's Guide*, Chicago, IL, ISBN 1-56827-333-9.
- Wainer, Howard. (1976) “Estimating Coefficients in Linear Models: It Don't Make No Nevermind”, *Psychological Bulletin*, Vol. 83, pp. 213-217.

- Zeller, Manfred. (2004) “Review of Poverty Assessment Tools”, Accelerated Microenterprise Advancement Project,
<http://www.povertytools.org/documents/Review%20of%20Poverty%20Assessment%20Tools.pdf>.
- Zeller, Manfred; and Gabriela Alcaraz V. (2005a) “Developing and Testing Poverty-Assessment Tools: Results from Accuracy Tests in Kazakhstan”, Accelerated Microenterprise Advancement Project,
<http://www.povertytools.org/documents/Kazakhstan%20Accuracy%20Report.pdf>.
- Zeller, Manfred; and Gabriela Alcaraz V. (2005b) “Developing and Testing Poverty-Assessment Tools: Results from Accuracy Tests in Uganda”, Accelerated Microenterprise Advancement Project,
<http://www.povertytools.org/documents/Uganda%20Accuracy%20Report.pdf>.
- Zeller, Manfred; Alcaraz V., Gabriela; and Julia Johannsen. (2005) “Developing and Testing Poverty-Assessment Tools: Results from Accuracy Tests in Peru”, Accelerated Microenterprise Advancement Project,
<http://www.povertytools.org/documents/Peru%20Accuracy%20Report.pdf>.
- Zeller, Manfred; Alcaraz V., Gabriela; and Julia Johannsen. (2004) “Developing and Testing Poverty-Assessment Tools: Results from Accuracy Tests in Bangladesh”, Accelerated Microenterprise Advancement Project,
<http://www.povertytools.org/documents/Bangladesh%20Accuracy%20Report.pdf>.