Simple Poverty Scorecard® Tool
Djibouti

Mark Schreiner and Jean Paul Sossou

16 November 2023

Ce document en Français et **l'application de collecte des données** sont disponibles sur **scorocs.com**
Kaadhkan dhibcaha iyo **appka-ururinta xogta** oo ku qoran afka soomaliga waxaad ka heli kartaa halkan **scorocs.com**
This document and a **data-collection app** are in English at **scorocs.com**

The Scorocs Simple Poverty Scorecard-brand poverty-assessment tool is a low-cost, transparent way for pro-poor programs in Djibouti to get to know their participants better so as to prove and improve their social performance. Responses to the scorecard's 10 questions can be used to:

- Assess poverty rates and numbers of poor people among in-coming participants
- Track changes in poverty among on-going participants
- Segment participants for differentiated treatment based on poverty

Acknowledgements

Data are from Djibouti's Direction de la Statistique et des Études Démographiques. Thanks go to Sekou Tidiani Konate, Omar Abdoulkader M., and Idriss Ali Soultan.

Scorocs® Simple Poverty Scorecard® Tool

			Full name	Identifier
Interview ID:				
Interview date:		Participant of record:		
Country:	DJI	Service agent:		
Scorecard:	001	Service point:		
Sampling weight:		Number of household members:		

Question	Response	Points
1. In which region or arrondissement does the household live? (*from enumerator knowledge*)	A. Tadjourah	0
	B. Dikhil	5
	C. Ali Sabieh	8
	D. Djibouti-ville (2nd or 4th arrondissement)	9
	E. Obock, or Djibouti-ville (3rd ou 5th arrondissement)	11
	F. Arta, or Djibouti-ville (1st arrondissement)	14
2. How many members does the household have? (*from **Back-page Worksheet***)	A. Eight or more	0
	B. Seven	6
	C. Six	10
	D. Five	15
	E. Four	20
	F. Three	24
	G. Two	31
	H. One	40
3. In the last 7 days, did any household member 5-years-old or older work at least one hour for a business or in self-employment, produce a good or service for another household or to make money for his/her own household, or work as a paid or unpaid intern? (*from **Back-page Worksheet***)	A. No	0
	B. Yes	5
4. Can the male head (or the husband of the female head) read and write?	A. No	0
	B. No male head (and the female head has no husband in the household)	2
	C. Yes	4
5. How many rooms are used by the household? (Exclude hallways, bathrooms, closets, storerooms, and rooms used only for business purposes)	A. One	0
	B. Two	3
	C. Three or more	8
6. What is the main material of the floor? (*as observed by enumerator*)	A. Dirt, or other	0
	B. Cement	3
	C. Wood, or tile	7
7. What is the household's main source of water?	A. River, creek, rainwater, dam, reservoir, underground cistern, unimproved well, improved well (without pump), borehole (well with pump), water truck, public standpipe, or other	0
	B. Piped by ONEAD to the residence's yard, or piped from a borehole	5
	C. Piped by ONEAD inside of the residence	7
8. Does the household have a TV?	A. No	0
	B. Yes	3
9. Is the main source of lighting for the household electricity from EDD?	A. No	0
	B. Yes	3
10. How many cell phones or radio phones does the household have?	A. None	0
	B. One	5
	C. Two or more	9

Score:

Back-page Worksheet

Members of the Household, Age, and Work Status

Fill out the scorecard header first. Include the interview's unique identifier (if known), the interview date, and the sampling weight of the participating household (if known). Then record the full name and unique identification number of the participant of record (who may differ from the respondent), of the service agent of the participant of record (who may differ from you the enumerator), and of the service point that the participant of record uses (if any and if known). Circle the response to the first scorecard question based on what you know about the region or arrrondissement where the household lives, without asking the respondent.

Then read to the respondent: *Please tell me the first name (or nickname) and age of each household member, starting with the head and his/her (eldest) spouse/conjugal partner (if there is one). A* household *is a group of people (regardless of blood or marital relationships) who recognize one member as the head of their household. Members of the household share all or part of their income and share expenses. Usually, members of a household live together in a residence and eat at least one meal together a day. The residence may be part of one building, all of one building, or more than one building. One person who lives alone (unmarried, widowed, or divorced) independently with his/her own income is the head of a one-person household. A household member must have lived with the household for more than six months as of the date of the interview or currently live with the household and plan to remain for a total duration of at least six months.*

Write down the name and age of each member, first for the head and then for his/her spouse (if there is one). There is no need to insist on the exact age unless it might be close to five. Record the sex of the head and the sex of his/her spouse (if there is one). For each member 5-years-old or older, ask: "In the past 7 days, did [NAME] work at least one hour . . .?", and record the response.

After you finish with all household members, record the number of members in the scorecard header next to "Number of household members:". Then circle the response to the second scorecard question. Record the response to the third scorecard question according to whether any household members 5-years-old or older work.

Read aloud the next two questions about the literacy of the male head and the number of rooms.

If you can observe the response to the sixth scorecard question about the main material of the floor, then record it without asking the respondent. Ask the respondent only if you cannot observe it yourself with complete certainty.

Finally, read aloud the last four questions.

Always keep in mind and apply the detailed instructions in the **Interview Guide**.

First name or nickname?	**Head or spouse of head?**	**How old is [NAME]?**	**If [NAME] is at least 5-years-old, then ask: "In the past 7 days, did [NAME] work at least one hour for a business or in self-employment, produce a good or service for another household or to make money for his/her own household, or work as a paid or unpaid intern?"**		
1.	Head (male) Head (female)		<5 years	No	Yes
2.	Spouse (female) Spouse (male) Other member		<5 years	No	Yes
3.	Other		<5 years	No	Yes
4.	Other		<5 years	No	Yes
5.	Other		<5 years	No	Yes
6.	Other		<5 years	No	Yes
7.	Other		<5 years	No	Yes
8.	Other		<5 years	No	Yes
9.	Other		<5 years	No	Yes
10.	Other		<5 years	No	Yes
11.	Other		<5 years	No	Yes
12.	Other		<5 years	No	Yes
Number of household members:	—	—	**Does anyone work?**	No	Yes

Figure 1: Conversion of scores to poverty likelihoods

	Poverty likelihood (%)															
	National					Intl. 2011 PPP				Percentile-based lines						
Score	Food	Minimum	100%	150%	200%	$1.90	$3.20	$5.50	$21.70	10th	20th	40th	50th	60th	80th	90th
0–24	64.2	84.3	91.3	97.5	99.1	78.8	93.5	99.1	99.9	64.5	80.5	93.5	97.1	98.1	99.1	99.7
25–28	39.5	61.1	83.2	96.5	97.8	55.9	91.1	97.9	99.9	36.4	62.5	91.1	95.0	96.2	99.1	99.7
29–32	29.3	51.9	81.1	95.8	97.4	43.3	86.8	97.6	99.9	24.9	53.7	87.5	92.9	95.6	99.1	99.7
33–35	18.3	39.5	63.4	90.2	96.2	32.8	74.2	96.0	99.9	17.7	41.0	74.3	86.9	92.1	99.1	99.7
36–38	14.0	32.3	53.4	86.6	95.0	26.6	60.8	92.3	99.9	12.4	31.6	60.9	75.4	87.5	99.1	99.7
39–40	6.1	20.7	50.5	82.7	94.6	14.1	59.0	90.8	99.9	5.7	21.1	59.5	71.9	85.2	97.0	99.6
41–42	4.7	20.7	49.4	77.4	94.1	14.1	59.0	88.2	99.9	3.7	20.5	59.5	71.9	80.7	95.6	99.4
43–44	2.5	18.4	37.0	68.0	88.0	10.8	43.9	86.3	99.9	2.1	16.2	43.9	58.1	77.3	94.1	97.8
45–46	2.3	11.8	36.1	66.7	88.0	3.9	43.9	85.9	99.9	2.1	9.0	43.9	51.7	71.7	94.1	97.8
47–48	1.6	10.1	29.5	64.3	87.8	3.5	36.3	83.3	99.9	1.7	6.9	36.3	49.4	67.5	94.1	97.8
49–50	0.4	4.7	22.5	64.3	85.4	2.8	27.8	79.8	99.9	0.4	3.4	27.8	49.4	65.6	91.6	97.8
51–52	0.4	2.2	13.5	46.2	78.0	1.7	15.7	72.4	99.9	0.2	2.0	15.7	34.8	49.0	86.5	96.0
53–54	0.3	1.4	6.2	41.1	71.4	0.4	8.4	60.2	99.8	0.1	0.7	8.4	21.1	41.1	82.9	96.0
55–56	0.3	1.1	4.0	33.2	62.5	0.4	4.3	53.3	99.8	0.1	0.5	4.5	14.3	27.6	76.6	94.1
57–59	0.3	0.6	1.7	23.5	50.5	0.4	1.9	39.6	99.7	0.1	0.5	2.3	8.5	21.2	70.2	90.9
60–62	0.1	0.3	1.5	9.5	28.7	0.1	1.9	27.8	98.4	0.0	0.1	2.1	5.0	9.5	46.9	82.1
63–65	0.0	0.0	1.3	9.0	21.5	0.0	1.9	16.6	97.1	0.0	0.0	2.1	3.8	9.5	31.9	69.8
66–69	0.0	0.0	0.5	4.2	11.8	0.0	0.7	7.6	96.5	0.0	0.0	0.7	1.9	2.6	19.9	54.7
70–74	0.0	0.0	0.0	0.5	3.7	0.0	0.0	0.6	92.1	0.0	0.0	0.0	0.0	0.4	6.7	34.2
75–100	0.0	0.0	0.0	0.1	1.3	0.0	0.0	0.1	60.2	0.0	0.0	0.0	0.0	0.1	0.3	7.6

Figure 2: Estimation errors in head-count poverty rates in a time period, along with margins of error and the α factor for finding margins of error and sample sizes

	Poverty lines															
	National					Intl. 2011 PPP				Percentile-based lines						
	Food	Minimum	100%	150%	200%	$1.90	$3.20	$5.50	$21.70	10th	20th	40th	50th	60th	80th	90th
Estimation error	–3.5	–2.6	–1.5	–4.1	–2.8	–2.1	+0.2	–4.1	0.0	+0.7	–0.5	+0.2	–2.7	–4.1	–1.2	0.0
Margin of error	2.3	2.2	2.4	2.0	1.7	2.0	2.3	1.8	0.5	2.0	2.0	2.3	2.5	2.1	1.4	1.1
α factor	1.43	1.13	1.13	1.04	0.89	0.74	1.15	1.01	0.76	0.83	1.14	1.37	1.09	1.00	1.03	0.93

Estimation errors from the scorecard with 1,000 bootstrap samples of n = 16,384 households from the validation sample.
Estimation errors are average differences between estimates and observed values, in percentage points.
Margins of error are ± percentage points with 90-percent confidence for samples of n = 1,024.
The α factor is used to calculate margins of error and sample sizes.
α is an average across 1,000 bootstrap samples of n = 256, 512, 1,024, 2,048, 4,096, 8,192, and 16,384.

Table of Contents

Figures

Scorocs® Simple Poverty Scorecard® Tool Djibouti

1. Introduction

The Scorocs Simple Poverty Scorecard-brand poverty-assessment tool for Djibouti is a low-cost, transparent way for pro-poor programs to get know their participants better so as to prove and improve their social performance.

1.1 Questions addressed by the scorecard

To address the question of "How many poor people does our program attract?", the scorecard can take a snapshot in a single time period with a census or a sample of in-coming households to estimate both head-count poverty rates as well as the number of poor people.

To address the question of "How has poverty changed for on-going participants?", the scorecard can be applied across two time periods with samples from a given population of on-going participants to estimate both net annual changes in head-count poverty rates as well as net annual changes in the number of poor people.

The scorecard can also be used for targeting, that is, to segment participants for differentiated treatment based on poverty.

It is difficult and costly for pro-poor programs to address these questions with the traditional direct approach to poverty assessment via expenditure surveys. A case in point is the 2017 Djibouti Household Survey (EDAM, *l'Enquête Djiboutienne Auprès des Ménages*) by Djibouti's *Direction de la Statistique es des Ètudes Démographiques* (DISED). The 2017 EDAM has more than 40 pages and asks more than 600 top-level questions, many of which have several follow-up questions or are repeated (for example, for each household member, each expenditure item, each consumer durable, or each daily meal).

1.2 How the scorecard works

The scorecard has 10 factual questions that are drawn from the exhaustive 2017 EDAM. Examples include: "How many rooms are used by the household?" and "Does the household have a TV?".

The 10 questions are selected to be:

- Inexpensive to collect, easy to answer quickly, and straightforward to verify
- Strongly and intuitively linked with poverty
- Liable to change over time as poverty changes
- Applicable in all regions and arrondissements of Djibouti

Each question has multiple-choice response options, with points assigned to each response. The points are zeroes or positive whole numbers. The points are derived from the statistical links between responses and expenditure-based poverty in the 2017 EDAM.

Adding up the points for a given household gives a *score* that ranges from 0 to 100. The lower the score, the poorer the household.

An enumerator can interview a household, record its responses on paper or **on a hand-held device**, and add up the household's score (if needed for on-the-spot segmentation) in about ten minutes.[1]

Back at the office or in the cloud, a household's score is converted into an estimated probability (the *poverty likelihood*) that the household is poor for a given poverty line. The links between scores and poverty likelihoods are based on EDAM data.

The average of poverty likelihoods across the members of sampled households is an estimate of the head-count poverty rate among people in the sampled population.

This estimated poverty rate may then be used to estimate:

- The number of poor people in in-coming households in a single time period
- The change in the net number of poor people in households of on-going participants across two time periods

[1] Responses on paper are entered in a spreadsheet or database later at an office.

1.3 Targeting

The scorecard can also be used to segment participating households for differentiated services. Unlike some other targeting tools—such as the World Bank's "proxy-means tests"[2]—the scorecard is transparent, freely available,[3] and tailored to the capabilities and purposes not of national governments but rather of local pro-poor programs. The feasible poverty-assessment tools for such programs are typically blunt (such as rules based on land ownership or housing quality) or subjective and relative (such as community-based, participatory wealth ranking facilitated by skilled field workers). Poverty assessments based on these approaches may be costly, their accuracy is unknown, and they are not comparable across places, programs, nor time.

1.4 Expenditure-based poverty

Djibouti's scorecard is a quantitative way to assess whether a program's participants have consumption expenditure below any of 16 poverty lines. The most-relevant line is Djibouti's "general poverty line" (called here the "100% of the national line") of about DJF306 per adult equivalent per day, giving a country-wide head-count poverty rate of 35.8 percent in 2017

A program uses only the poverty line(s) that fit its context and mission. For example, a program may report poverty estimates to funders based a national line while internally using a percentile-based line.

1.5 Transparency

The scorecard's design aims to make its workings clear to program managers. The tool's adoption stems from the low cost of its short interviews and from the fact that managers can see for themselves how the scorecard works and that its approach makes sense. Similar tools have been around for decades, but pro-poor programs have rarely used them. This is not because these tools are inaccurate, but because *how* they work is unclear or hidden.

---

[2] **Coady, Grosh, and Hoddinott**, 2004.

[3] Djibouti's scorecard is not in the public domain; it is copyright © 2023 Scorocs.

When scorecard projects fail, the cause is not usually inaccuracy but rather a program's failure to commit to the work-a-day project management needed to integrate the scorecard in the program's processes and to train and convince employees to use the tool properly.[4] For tool-based estimates of social outcomes such as poverty, data scientists have long known that there is almost no trade-off between the straightforward and transparent versus the complex and opaque.[5] Project risk is less technical and more human, not statistics but organizational-change management.

1.6 Assumptions and estimation errors

Like all predictive tools, the scorecard makes two fundamental assumptions:

- The scored sample is representative of the same population as that whose data was used to construct the scorecard
- The links between responses and poverty are the same in the scored sample as in the population whose data was used to construct the scorecard

Of course, these assumptions do not hold to some unknown degree.[6] In particular:

- A given program's participants are not representative of Djibouti overall
- Over time, the links between responses and poverty drift or shift

Scorecard estimates have errors because the scorecard incorrectly acts as if the links between responses and poverty in all scored samples and in all time periods are the same as in the construction sample from the 2017 EDAM. Reality diverges further from assumptions as:

- More time passes since the collection of construction data
- A program's participants differ from the country's general population
- Attrition has changed the composition of a cohort of on-going participants
- Change has been rapid (say, due to war, plague, or changes in the program itself)[7]

---

[4] **Schreiner**, 2002.

[5] **Dupriez**, 2018; **Caire and Schreiner**, 2012; **Schreiner**, 2012; **Hand**, 2006; **Lovie and Lovie**, 1986; **Kolesar and Showers**, 1985; **Stillwell, Barron, and Edwards**, 1983; **Dawes**, 1979; **Wainer**, 1976; **Myers and Forgy**, 1963.

[6] **Diamond *et al.***, 2016; **Tarozzi and Deaton**, 2009.

[7] For example, the 2020/23 economic downturn due to COVID−19 changed the links between poverty and questions, but the Djibouti scorecard still uses 2017 links.

For any particular scorecard and scored sample, the estimation error due to migration away from the assumptions is unknown. It is known, however, that the scorecard's targeting is robust. That is, the extent to which assumptions diverge from reality is not strongly linked with the extent to which the scorecard gives lower scores to more-poor households and higher scores to less-poor households. It is also known that the scorecard's estimation errors are larger when estimating changes in poverty across two periods (or across two scorecards) than when estimating poverty in one period or across two periods with a single scorecard.

There are no rules nor formulas that automatically signal when estimation error is too large for estimates to be useful. Program managers must make their own judgments based on common sense and on what they know about their context and their participants from non-scorecard sources.

In practice, scorecard estimates often serve as a basic check on whether a pro-poor program is indeed *pro-poor*. The estimates address existential questions such as:

- "How many in-coming participants are below the national poverty line?"
- "Are in-coming participants poorer than the average person in the area where we work?"
- "Are our poor participants more likely to rise above a poverty line than the average poor person in the area where we work?"

For such existential checks on whether a program lives out its purported social mission, estimation errors will often be small enough to be immaterial.

1.7 Estimation errors when assumptions hold

If the scorecard's assumptions do hold, then the scorecard estimators are statistically *unbiased*. That is, the true value in the population matches the average of scorecard estimates from repeated samples.

The assumptions do hold when the scorecard is tested against households in the validation sample from the 2017 EDAM that are not used to construct the scorecard. Smaller errors in this ideal case imply smaller-than-otherwise errors in real-world use.

Even so, there are estimation errors on average in the validation sample because there is only one scorecard, and it is derived from one construction sample and applied to a single validation sample. **Figure 2** documents the error for estimates of poverty rates in one time period, allowing scorecard users to adjust for this error.

1.8 The rest of this document

Section 2: How to collect responses to scorecard questions from households and then convert those responses to poverty likelihoods

Section 3: **How to calculate scorecard estimates**

- Estimates in a single time period of:
  - **Head-count poverty rates in a single time period**
  - **Number of poor people in a single time period**
- Estimates across two time periods of:
  - **Annual net change in poverty rates with one sample scored twice**
  - **Annual net change in the number of poor people with one sample scored twice**
  - **Annual net change in poverty rates with two independent samples**
  - **Annual net change in the number of poor people with two independent samples**

Section 4: **How to design scorecard surveys and samples**

Section 5: **How to use scores for targeting**

After **Section 5**, the **Interview Guide** tells how to ask questions—and how to interpret responses—so as to mimic practice in Djibouti's 2017 EDAM as closely as possible. The **Interview Guide** (and the **Back-page Worksheet**) are integral parts of the scorecard. Do not ignore them.

The annexes provide details for advanced users:

Annex 1: **Data used for construction and validation**

Annex 2: **Definition of poverty**

Annex 3: **Scorecard construction**

Annex 4: **Estimates of poverty likelihoods**

Annex 5 : **Error and margins of error**

Annex 6: **Formulas for sample size**

Details on cited **References** appear at the end.

2. How to collect responses to scorecard questions from households and then convert those responses to poverty likelihoods

This section tells how to:

- Collect a household's responses to scorecard questions
- Convert responses to points
- Add up points to get scores
- Convert scores to poverty likelihoods

The next section tells how to combine poverty likelihoods from a sample of households to estimate poverty.

2.1 Instructions for enumerators

An *enumerator* asks a scorecard's questions to a respondent and then records the responses. An enumerator may or may not be same as the program's service agent (if any) who is associated with a participating household.

Enumerators should interview a sampled household at the household's residence using an **app on a hand-held device** or a paper scorecard along with the **Back-page Worksheet**. Following the **Interview Guide**, enumerators should:

- Record administrative information in the scorecard header:
  - Interview identifier (if known)
  - Interview date (required)
  - Country code ("DJI", pre-filled)
  - Scorecard code ("001", pre-filled)
  - Sampling weight assigned to the household by the survey design (if any and if known)
- Record names and identifiers (if known) in the scorecard header:
  - *Participant of record*. This is the member of the household whose identifying information is recorded on-file with the pro-poor program. Often, the participant of record is the adult member of the household who interacts directly with the program. He/she may or may not be the same as the respondent who responds to the scorecard questions. For example, a participant of record for a microfinance program is often a borrower or a saver, and a participant of record with a child-health program might be a child or a child's parent or guardian

  - *Service agent* (if there is one and if known). This is the participant of record's main, on-going point of contact with the program. The service agent may or may not be the same as the enumerator. For example, the service agent in a microfinance program is often a loan officer or savings collector, and the service agent in a child-health program might be a community health-care worker or a nurse practitioner
  - *Service point* (if there is one). This is the program office that is relevant to the participant of record. The service point is usually the base of operations for the service agent (if there is one) who serves the participant of record or where the participant of record usually does program business. For example, the service point for a microfinance program is often a branch, and the service point for a child-health program might be a community health post
- Mark the response to the first scorecard question ("In what region or arrondissement does the household live?"). If the enumerator already knows the region or arrondissement (as is usually the case), then the question need not be asked directly of the respondent
- Complete the **Back-page Worksheet** with each household member's first name (or nickname), age, and work status
- If using a paper scorecard, then use the **Back-page Worksheet** to record:
  - The number of household members in the header next to "Number of household members:"
  - The response to the second scorecard question ("How many members does the household have?")
  - The response to the third scorecard question ("In the last 7 days, did any household member 5-years-old or older work at least one hour?")
- Read aloud the fourth and fifth questions about the literacy of the male head and the number of rooms one-by-one and in order, marking the responses given by the respondent
- Do not read the response options for a scorecard question to the respondent
- If possible, the enumerator should not read aloud the sixth scorecard question ("What is the main material of the floor?"). Instead, the enumerator should observe and record the main material of the floor without asking the question directly of the respondent. However, if the enumerator is not completely certain about the main material of the floor, then he/she should ask the question directly of the respondent
- Read aloud the remaining four questions one-by-one and in order, marking the responses given by the respondent

- When marking a response on paper, write each point value in the far right-hand column. Then make single circle around the pre-printed response, the pre-printed points, and the hand-written points. This helps to reduce later data-entry mistakes
- Add up the points to get the score (if needed on-the-spot and if using a paper scorecard)
- Implement targeting policy (if any) based on the score
- Upload the data with a **mobile data-collection tool**, or deliver the filled-out paper scorecard to a central office for data entry, reporting, and analysis

2.2 Header, Back-page Worksheet, Interview Guide, and audits

Fill out the scorecard header as best you can; do not skip it. Scorecard estimates are more useful if they can be linked—via names or identifiers—to a program's existing data on the participant of record, service agent, or service point. Record the types of identifiers that are used in the program's databases, be they program-specific or government-issued. Be sure to record the number of household members not only indirectly via the scorecard's second question but also directly in the header.

Do not leave fields in the header blank. If the information does not exist, is not applicable, or is unknown, then write "NONE" or "UNKNOWN".

Likewise, do not skip the **Back-page Worksheet**. Take the time to read the definition of *household* to the respondent and to fill out the roster member-by-member. If you cut corners, many respondents will miscount or apply the wrong definition of *household*. Completing the **Back-page Worksheet** improves data quality because it mimics the practice of Djibouti's DISED in the 2017 EDAM. The accuracy of the scorecard's estimates depends on the quality of recorded responses and especially strongly on an accurate count of household members. Working through the **Back-page Worksheet** gives the best data.

Throughout the interview, apply the instructions in the **Interview Guide**. Enumerators must be thoroughly trained on the **Interview Guide** before they do any interviews, and they should carry a copy of the **Interview Guide** with them to each interview.[8] Even though the scorecard is less difficult than other poverty-assessment tools, training and explicit definitions of the scorecard's terms and concepts are still essential.[9] Enumerators must scrupulously study and follow the **Interview Guide**.

Finally, on-going quality-control audits are wise if a program or its service agents collect their own data and if they believe that they have an incentive to exaggerate poverty estimates (for example, if they expect to be rewarded for higher poverty rates).[10]

[8] The **Interview Guide** is the only source of guidance for enumerators. All other issues of interpretation should be left to the judgment of enumerators and respondents, as this seems to be what Djibouti's DISED did in the 2017 EDAM.
[9] Merely reading through the scorecard with enumerators is not adequate training.
[10] **Matul and Kline**, 2003. If a program does not want enumerators or respondents to know the scorecard's points, then it can use a **mobile data-collection app** or a paper version of the scorecard that omits the points, with scores computed later at an office. Even if points are hidden, however, enumerators and respondents can use common sense to guess how responses are linked with poverty.

Figure 3: First example household, filled-in scorecard

			Name	Identifier
Interview ID:	A123			
Interview date:	13JUN2021	Participant of record:	ANNA JACKSON	1V0276FZ7
Country:	DJI	Service agent:	UNKNOWN	UNKNOWN
Scorecard:	001	Service point:	NORTHWEST CLINIC	NWC
Sampling weight:	UNKNOWN	Number of household members:	Eight	

Question	Response	Points	
1. In which region or arrondissement does the household live? (*from enumerator knowledge*)	A. Tadjourah	0	
	B. Dikhil	**5**	**5**
	C. Ali Sabieh	8	
	D. Djibouti-ville (2nd or 4th arrondissement)	9	
	E. Obock, or Djibouti-ville (3rd ou 5th arrondissement)	11	
	F. Arta, or Djibouti-ville (1st arrondissement)	14	
2. How many members does the household have? (*from **Back-page Worksheet***)	**A. Eight or more**	**0**	**0**
	B. Seven	6	
	C. Six	10	
	D. Five	15	
	E. Four	20	
	F. Three	24	
	G. Two	31	
	H. One	40	
3. In the last 7 days, did any household member 5-years-old or older work at least one hour for a business or in self-employment, produce a good or service for another household or to make money for his/her own household, or work as a paid or unpaid intern?	A. No	0	
	B. Yes	**5**	**5**
4. Can the male head (or the husband of the female head) read and write?	**A. No**	**0**	**0**
	B. No male head (and the female head has no husband in household)	2	
	C. Yes	4	
5. How many rooms are used by the household? (Exclude hallways, bathrooms, closets, storerooms, and rooms used only for business purposes)	A. One	0	
	B. Two	**3**	**3**
	C. Three or more	8	
6. What is the main material of the floor? (*as observed by enumerator*)	A. Dirt, or other	0	
	B. Cement	**3**	**3**
	C. Wood, or tile	7	
7. What is the household's main source of water?	**A. River, creek, rainwater, dam, reservoir, underground cistern, unimproved well, improved well (without pump), borehole (well with pump), water truck, public standpipe, or other**	**0**	**0**
	B. Piped by ONEAD to the residence's yard, or piped from a borehole	5	
	C. Piped by ONEAD inside of the residence	7	
8. Does the household have a TV?	A. No	0	
	B. Yes	**3**	**3**
9. Is the main source of lighting for the household electricity from EDD?	A. No	0	
	B. Yes	**3**	**3**
10. How many cell phones or radio phones does the household have?	**A. None**	**0**	**0**
	B. One	5	
	C. Two or more	9	

scorocs.com

Score: 5 + 0 + 5 + 0 + 3 + 3 + 0 + 3 + 3 + 0 = 22

Figure 4: First example household, filled-in Back-page Worksheet

First name or nickname?	Head or spouse of head?	How old is [NAME]?	If [NAME] is at least 5-years-old, then ask: "In the past 7 days, did [NAME] work at least one hour for a business or in self-employment, produce a good or service for another household or to make money for his/her own household, or work as a paid or unpaid intern?"		
1. **BILLY**	Head (male) Head (female)	**38**	<5 years	No	Yes
2. **ANNA**	Spouse (female) Spouse (male) Other	**40**	<5 years	No	Yes
3. **CHARLES**	Other	**24**	<5 years	No	Yes
4. **DARLA**	Other	**21**	<5 years	No	Yes
5. **EUGENE**	Other	**18**	<5 years	No	Yes
6. **FANNY**	Other	**10**	<5 years	No	Yes
7. **GERTRUDE**	Other	**4**	<5 years	No	Yes
8. **HENRY**	Other	**2**	<5 years	No	Yes
9.	Other		<5 years	No	Yes
10.	Other		<5 years	No	Yes
11.	Other		<5 years	No	Yes
12.	Other		<5 years	No	Yes
Number of household members: EIGHT		**—**	**Does anyone work?**	**No**	**Yes**

2.3 First example household

The points for the first example household's responses add up to a score of 22 (**Figure 3** and **Figure 4**).

For a given poverty line, **Figure 1** lists poverty likelihoods by score range. A score of 22 falls in the first range of 0−24. For 100% of the national poverty line, the poverty likelihood for scores of 0−24 is 91.3 percent. That is, the scorecard estimates that 91.3 percent of households in Djibouti with a score of 0−24 have consumption expenditure below 100% of the national line.

Figure 5: The first example household's score of 22 corresponds with a poverty likelihood of 91.3 percent for 100% of the national line (excerpted from Figure 1)

		National			
Score	**Food**	**Minimum**	**100%**	**150%**	**200%**
0−24	64.2	84.3	91.3	97.5	99.1
25−28	39.5	61.1	83.2	96.5	97.8
29−32	29.3	51.9	81.1	95.8	97.4
33−35	18.3	39.5	63.4	90.2	96.2
36−38	14.0	32.3	53.4	86.6	95.0
39−40	6.1	20.7	50.5	82.7	94.6
41−42	4.7	20.7	49.4	77.4	94.1
43−44	2.5	18.4	37.0	68.0	88.0
45−46	2.3	11.8	36.1	66.7	88.0
. . .	. . .	. . .	. . .	. . .	. . .

Figure 6: Second example household, filled-in scorecard

			Name	Identifier
Interview ID:	B456			
Interview date:	30JUN2021	Participant of record:	ALICE BROWN	2W3120ZG8
Country:	DJI	Service agent:	UNKNOWN	UNKNOWN
Scorecard:	001	Service point:	NORTHWEST CLINIC	NWC
Sampling weight:	UNKNOWN	Number of household members:	FIVE	

Question	Response		Points	
1. In which region or arrondissement does the household live? (*from enumerator knowledge*)	A. Tadjourah		0	
	B. Dikhil		**5**	**5**
	C. Ali Sabieh		8	
	D. Djibouti-ville (2nd or 4th arrondissement)		9	
	E. Obock, or Djibouti-ville (3rd ou 5th arrrondissement)		11	
	F. Arta, or Djibouti-ville (1st arrondissement)		14	
2. How many members does the household have? (*from **Back-page Worksheet***)	A. Eight or more		0	
	B. Seven		6	
	C. Six		10	
	D. Five		**15**	**15**
	E. Four		20	
	F. Three		24	
	G. Two		31	
	H. One		40	
3. In the last 7 days, did any household member 5-years-old or older work at least one hour for a business or in self-employment, produce a good or service for another household or to make money for his/her own household, or work as a paid or unpaid intern?	**A. No**		**0**	**0**
	B. Yes		5	
4. Can the male head (or the husband of the female head) read and write?	A. No		0	
	B. No male head (and the female head has no husband in household)		**2**	**2**
	C. Yes		4	
5. How many rooms are used by the household? (Exclude hallways, bathrooms, closets, storerooms, and rooms used only for business purposes)	**A. One**		**0**	**0**
	B. Two		3	
	C. Three or more		8	
6. What is the main material of the floor? (*as observed by enumerator*)	**A. Dirt, or other**		**0**	**0**
	B. Cement		3	
	C. Wood, or tile		7	
7. What is the household's main source of water?	A. River, creek, rainwater, dam, reservoir, underground cistern, unimproved well, improved well (without pump), borehole (well with pump), water truck, public standpipe, or other		0	
	B. Piped by ONEAD to the residence's yard, or piped from a borehole		**5**	**5**
	C. Piped by ONEAD inside of the residence		7	
8. Does the household have a TV?	**A. No**		**0**	**0**
	B. Yes		3	
9. Is the main source of lighting for the household electricity from EDD?	**A. No**		**0**	**0**
	B. Yes		3	
10. How many cell phones or radio phones does the household have?	A. None		0	
	B. One		**5**	**5**
	C. Two or more		9	

scorocs.com

Score: 5 + 15 + 0 + 2 + 0 + 0 + 5 + 0 + 0 + 5 = 32

Figure 7: Second example household, filled-in Back-page Worksheet

First name or nickname?	Head or spouse of head?	How old is [NAME]?	If [NAME] is at least 5-years-old, then ask: "In the past 7 days, did [NAME] work at least one hour for a business or in self-employment, produce a good or service for another household or to make money for his/her own household, or work as an paid or unpaid intern?"		
1. ALICE	Head (male) Head (female)	29	<5 years	No	Yes
2. BERNITA	Spouse (female) Spouse (male) Other	11	<5 years	No	Yes
3. CARLOS	Other	9	<5 years	No	Yes
4. DARLENE	Other	5	<5 years	No	Yes
5. EVELYN	Other	3	<5 years	No	Yes
6.	Other		<5 years	No	Yes
7.	Other		<5 years	No	Yes
8.	Other		<5 years	No	Yes
9.	Other		<5 years	No	Yes
10.	Other		<5 years	No	Yes
11.	Other		<5 years	No	Yes
12.	Other		<5 years	No	Yes
Number of household members: FIVE		—	Does anyone work?	No	Yes

2.4 Second example household

The points for the second example household's responses add up to a score of 32 (**Figure 6** and **Figure 7**).

In **Figure 1**, a score of 32 falls in the range of 29−32. For 100% of the national poverty line, the poverty likelihood for scores of 29−32 is 81.1 percent. The scorecard estimates that 81.1 percent of households in Djibouti with a score of 29−32 have consumption expenditure below 100% of the national line.

Figure 8: The second example household's score of 32 coresponds with a poverty likelihood of 81.1 percent for 100% of the national line (excerpt from Figure 1)

		National			
Score	**Food**	**Minimum**	**100%**	**150%**	**200%**
0−24	64.2	84.3	91.3	97.5	99.1
25−28	39.5	61.1	83.2	96.5	97.8
29−32	29.3	51.9	81.1	95.8	97.4
33−35	18.3	39.5	63.4	90.2	96.2
36−38	14.0	32.3	53.4	86.6	95.0
39−40	6.1	20.7	50.5	82.7	94.6
41−42	4.7	20.7	49.4	77.4	94.1
43−44	2.5	18.4	37.0	68.0	88.0
45−46	2.3	11.8	36.1	66.7	88.0
. . .	. . .	. . .	. . .	. . .	. . .

3. How to calculate scorecard estimates

This section tells how to estimate:

- Head-count poverty rates for a single time period for in-coming participants
- Net changes in poverty rates across two time periods for on-going participants

It also tells how to use these estimated poverty rates to estimate:

- Number of poor people in the households of in-coming participants
- Net change in the number of poor people in the households of on-going participants

3.1 Head-count poverty rates in a single time period

The *head-count poverty rate* is the share of people in participating households in which total household consumption expenditure (divided by the number of members in the household or by the number of adult equivalents in the household) is below a given poverty line.

An estimate of the head-count poverty rate is the household-size-weighted average of poverty likelihoods from a scored sample, adjusted for the scorecard's known estimation error.

To illustrate the calculation, suppose that a pro-poor program opens a new service point in rural Dikhil in 2021. In that calendar year, it enrolls 1,000 in-coming households, from which it scores a simple random sample[11] of two households.[12]

The program judges that 100% of the national poverty line is the most-relevant line for its purposes. For that line and for estimates of poverty rates in one period, the scorecard's known estimation error is −1.5 percentage points (**Figure 2**).

The first example household has eight members and is interviewed on June 13, 2021 (**Figure 3** and **Figure 4**). With a score of 22, it has a poverty likelihood for 100% of the national line of 91.3 percent (**Figure 1**).

The second example household has five members and is interviewed on June 30, 2021 (**Figure 6** and **Figure 7**). Its score of 32 corresponds with a poverty likelihood of 81.1 percent.

---

[11] In a *simple random sample*, all households in the population have the same selection probability. This paper does not discuss samples in which different households have different selection probabilities.

[12] Of course, estimates based on such an unrealistically small sample have wide margins of error, but a small sample facilitates the arithmetic in the examples here.

The estimated head-count poverty rate for the population of in-coming households in the 2021 calendar-year cohort in this new rural Dikhil service point is the household-size-weighted average of the estimated poverty likelihoods of the sampled households, less the known estimation error. Expressing poverty likelihoods and the estimation error as proportions between 0 and 1 rather than percentages between 0 and 100, this is:

$$\frac{8 \cdot 0.913 + 5 \cdot 0.811}{8 + 5} - (-0.015) \approx \frac{11.36}{13} + 0.015 \approx 0.889 = 88.9 \,\text{percent}.$$

The eight in the "8 · 0.913" term is the number of members (household size) in the first household, and 0.913 is the first household's estimated poverty likelihood as proportion.

In the same way, the five in "5 · 0,811" is the number of members in the second household, and 0.811 is the second household's estimated poverty likelihood.

The "8 + 5" is the sum of the weights—that is, the number of household members—across the two sampled households.

The "−0.015" is the scorecard's estimation error for this poverty line (**Figure 2**). Because unadjusted estimates tend to be too low by 1.5 percentage points, they are adjusted upwards by subtracting −1.5. This is akin to how an archer whose arrows tend to miss a little to the left of the bulls-eye will adjust his/her aim to be a little to the right of the bulls-eye.

The estimated head-count poverty rate for the population is 88.9 percent. Again, this is the household-size-weighted average of the two sampled households' poverty likelihoods, adjusted for the known estimation error.[13]

With hundreds or thousands of interviewed households, the calculations should be done by an **app** or in a spreadsheet, as modeled in **Figure 9** below.

---

[13] Be careful; the estimated poverty rate is *not* the single poverty likelihood associated with the household-size-weighted average score, which here is (8·22 + 5·32) ÷ (8 + 5) ≈ 26. This average score of 26 corresponds to a poverty likelihood of 83.2 percent (**Figure 1**), giving an error-adjusted poverty rate of 83.2 − (−1.5) = 84.7 percent. This differs from the 88.9 percent found as the household-size-weighted average of the two individual likelihoods associated with each of the two scores. Unlike likelihoods, scores are ordinal symbols, like colors in the spectrum or syllables in a solfège scale. Because scores are ordinal, they cannot be added up nor averaged. Only three operations are valid for scores: conversion to likelihoods, analysis of distributions, or comparison with a cut-off for segmentation (**Schreiner**, 2012). In general, programs should analyze likelihoods, not scores.

Figure 9: Spreadsheet calculation to estimate the head-count poverty rate and number of poor people in a population of in-coming participants in a period

	A	B	C	D	E	F	G
1	**Survey**	**Interview date**	**ID of direct participant**	**Number of household members**	**Score**	**Poverty likelihood (%)**	**Estimated number of poor household members**
2	Baseline	13-Jun-21	1V0276FZ7	8	22	91.3	7.30 = (D2*F2)/100
3	Baseline	30-Jun-21	2W3120ZG8	5	32	81.1	4.06 = (D3*F3)/100
4			**Sum:**	13 = SUM(D2:D3)			11.36 = SUM(G2:G3)
5			**Average:**	6.5 = AVERAGE(D2:D3)			
6							
7		**Estimated scorecard error for this poverty line (percentage points):**					–1.5
8							
9				**Estimated head-count poverty rate (%):**			88.9 = (G4/D4)*100-G7
10							
11				**Households in the population:**			1,000
12							
13				**People in households in the population:**			6,500 = G11*D5
14							
15				**Number of poor people in population:**			5,777 = (G9/100)*G13
16	Rows of data are sorted by Survey, then by Interview date, then by the ID of the participant of record.						

This estimate in a single time period tends to be more relevant for in-coming participants who joined a program in the current period than for on-going participants who joined in past periods. This is because fulfilling a pro-poor mission implies that some share of new participants be poor by some definition of *poverty*.[14] To be pro-poor, a bare-minimum standard is that the poverty rate of in-coming participants exceed that of the country as a whole or that of the area where the program works.

To help with benchmarking poverty-rate estimates, **Figure 10** reports head-count poverty rates from the 2017 EDAM for all 16 poverty lines by urban/rural/all for Djibouti overall and for each of its 10 regions and arrondissements. In the example of rural Dikhil, the head-count poverty rate for 100% of the national line is 82.0 percent. Thus, the example program is pro-poor in the sense that its in-coming participants have an above-average estimated poverty rate for that area (88.9 percent).

The text that illustrates the calculation of the scorecard estimate of the number of poor people in a single time period follows after **Figure 10**, which stretches across the next four pages. The regions and arrondissements in **Figure 10** begin with Djibouti overall, followed by the 10 regions and arrondissments in DISED's usual order for reporting.

[14] The Djibouti scorecard uses a definition of *poverty* based on consumption expenditure. Other common definitions of *poverty* include: being rural, agricultural, landless, or unemployed; living in a given area; having a head who is illiterate, female, or an ethnic minority; or having a member who is pregnant, handicapped, elderly, or young.

Figure 10: (Djibouti overall; Djibouti-Ville, 1st Arrondissement; and Djibouti-Ville, 2nd Arrondissement): Poverty lines and head-count poverty rates by urban/rural/all in 2017

Region/ Area	Line or Rate	*n*	Poverty lines and poverty rates															
			National					Intl. 2011 PPP				Percentile-based lines						
			Food	Minimum	100%	150%	200%	$1.90	$3.20	$5.50	$21.70	10th	20th	40th	50th	60th	80th	90th
Djibouti	**Urban**	Line	218	306	415	622	830	212	356	612	2,417	163	235	359	428	506	743	1,063
		Rate	5.6	13.7	28.1	53.6	69.8	9.8	32.1	65.7	98.1	4.2	12.2	32.2	43.0	54.2	77.0	88.3
		2,671																
	Rural	Line	218	306	415	622	830	212	356	612	2,417	163	235	359	428	506	743	1,063
		Rate	43.1	62.6	78.4	90.8	96.0	57.4	82.9	95.2	99.9	42.5	63.3	83.3	88.7	92.6	97.7	99.2
		1,803																
	All	Line	218	306	415	622	830	212	356	612	2,417	163	235	359	428	506	743	1,063
		Rate	11.3	21.1	35.8	59.2	73.8	17.0	39.8	70.2	98.3	10.0	20.0	40.0	50.0	60.0	80.1	89.9
		4,474																
Djibouti-ville, 1er	**Urban**	Line	218	306	415	622	830	212	356	612	2,417	163	235	359	428	506	743	1,063
		Rate	0.4	4.7	13.0	27.8	41.1	3.5	14.6	36.2	93.3	0.4	3.9	14.6	22.9	29.7	49.3	71.2
		389																
	Rural	Line	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
		Rate	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
		—																
	All	Line	218	306	415	622	830	212	356	612	2,417	163	235	359	428	506	743	1,063
		Rate	0.4	4.7	13.0	27.8	41.1	3.5	14.6	36.2	93.3	0.4	3.9	14.6	22.9	29.7	49.3	71.2
		389																
Djibouti-ville, 2ème	**Urban**	Line	218	306	415	622	830	212	356	612	2,417	163	235	359	428	506	743	1,063
		Rate	3.8	9.8	22.0	45.6	61.2	6.1	24.9	56.7	97.5	2.6	7.3	24.9	33.3	45.5	69.8	83.7
		388																
	Rural	Line	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
		Rate	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
		—																
	All	Line	218	306	415	622	830	212	356	612	2,417	163	235	359	428	506	743	1,063
		Rate	3.8	9.8	22.0	45.6	61.2	6.1	24.9	56.7	97.5	2.6	7.3	24.9	33.3	45.5	69.8	83.7
		388																

Source: 2017 EDAM.

Poverty rates are percentages.

National poverty lines are DJF per adult equivalent, per day.

International 2011 PPP lines and percentile-based lines are DJF per-person, per-day.

All poverty lines are DJF in prices in Djibouti as a whole on average in May 2017.

Figure 10: (Djibouti-Ville, 3ème Arrondissement; Djibouti-Ville, 4ème Arrondissement; and Djibouti-Ville, 5ème Arrondissement;): Poverty lines and head-count poverty rates by urban/rural/all in 2017

Region/ Area	Line or Rate	*n*	Poverty lines and poverty rates															
			National					Intl. 2011 PPP				Percentile-based lines						
			Food	Minimum	100%	150%	200%	$1.90	$3.20	$5.50	$21.70	10th	20th	40th	50th	60th	80th	90th
Djibouti-ville, 3ème	Urban	Line	218	306	415	622	830	212	356	612	2,417	163	235	359	428	506	743	1,063
		Rate	3.5	8.5	15.4	39.5	56.7	5.8	19.1	50.9	93.0	3.0	6.7	19.1	27.4	37.9	63.3	76.1
		n	397															
	Rural	Line	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
		Rate	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
		n	—															
	All	Line	218	306	415	622	830	212	356	612	2,417	163	235	359	428	506	743	1,063
		Rate	3.5	8.5	15.4	39.5	56.7	5.8	19.1	50.9	93.0	3.0	6.7	19.1	27.4	37.9	63.3	76.1
		n	397															
Djibouti-ville, 4ème	Urban	Line	218	306	415	622	830	212	356	612	2,417	163	235	359	428	506	743	1,063
		Rate	8.0	18.3	35.2	65.1	82.7	13.0	39.2	78.7	100.0	5.4	15.3	39.4	54.5	66.1	87.4	95.1
		n	425															
	Rural	Line	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
		Rate	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
		n	—															
	All	Line	218	306	415	622	830	212	356	612	2,417	163	235	359	428	506	743	1,063
		Rate	8.0	18.3	35.2	65.1	82.7	13.0	39.2	78.7	100.0	5.4	15.3	39.4	54.5	66.1	87.4	95.1
		n	425															
Djibuoti-ville, 5ème	Urban	Line	218	306	415	622	830	212	356	612	2,417	163	235	359	428	506	743	1,063
		Rate	5.8	15.8	33.6	59.0	75.6	11.1	38.7	71.8	98.6	4.7	15.6	39.0	48.3	59.1	83.6	91.8
		n	436															
	Rural	Line	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
		Rate	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
		n	—															
	All	Line	218	306	415	622	830	212	356	612	2,417	163	235	359	428	506	743	1,063
		Rate	5.8	15.8	33.6	59.0	75.6	11.1	38.7	71.8	98.6	4.7	15.6	39.0	48.3	59.1	83.6	91.8
		n	436															

Source: 2017 EDAM.

Poverty rates are percentages.

National poverty lines are DJF per adult equivalent, per day.

International 2011 PPP lines and percentile-based lines are DJF per-person, per-day.

All poverty lines are DJF in prices in Djibouti as a whole on average in May 2017.

Figure 10: (Ali Sabieh, Dikhil, and Tadjourah): Poverty lines and head-count poverty rates by urban/rural/all in 2017

Region/ Area	Line or Rate	*n*	Poverty lines and poverty rates															
			National					Intl. 2011 PPP				Percentile-based lines						
			Food	Minimum	100%	150%	200%	$1.90	$3.20	$5.50	$21.70	10th	20th	40th	50th	60th	80th	90th
Ali Sabieh	Urban	Line	218	306	415	622	830	212	356	612	2,417	163	235	359	428	506	743	1,063
		Rate	3.4	7.6	17.1	44.9	63.4	6.2	23.3	59.8	99.6	3.1	7.9	23.3	33.7	48.9	75.1	92.8
		n	230															
	Rural	Line	218	306	415	622	830	212	356	612	2,417	163	235	359	428	506	743	1,063
		Rate	32.4	56.6	77.4	89.9	95.9	48.4	82.6	94.7	100.0	35.3	55.9	82.9	87.5	92.2	97.7	99.6
		n	265															
	All	Line	218	306	415	622	830	212	356	612	2,417	163	235	359	428	506	743	1,063
		Rate	15.0	27.2	41.2	62.9	76.4	23.0	47.0	73.8	99.8	16.0	27.1	47.1	55.2	66.2	84.1	95.5
		n	495															
Dikhil	Urban	Line	218	306	415	622	830	212	356	612	2,417	163	235	359	428	506	743	1,063
		Rate	14.1	24.5	38.0	66.7	81.1	23.3	39.3	79.3	100.0	12.0	23.6	39.3	56.2	69.4	86.2	93.2
		n	133															
	Rural	Line	218	306	415	622	830	212	356	612	2,417	163	235	359	428	506	743	1,063
		Rate	48.7	68.7	82.0	93.6	97.7	66.1	85.6	97.7	100.0	48.6	69.7	86.2	91.8	95.7	99.1	99.7
		n	363															
	All	Line	218	306	415	622	830	212	356	612	2,417	163	235	359	428	506	743	1,063
		Rate	36.3	52.9	66.2	84.0	91.8	50.8	69.1	91.1	100.0	35.5	53.2	69.4	79.1	86.3	94.5	97.4
		n	496															
Tadjourah	Urban	Line	218	306	415	622	830	212	356	612	2,417	163	235	359	428	506	743	1,063
		Rate	13.5	21.7	42.2	73.1	81.2	18.1	53.0	78.9	100.0	13.5	26.9	53.0	61.1	71.9	85.7	92.9
		n	56															
	Rural	Line	218	306	415	622	830	212	356	612	2,417	163	235	359	428	506	743	1,063
		Rate	60.3	78.5	88.9	96.9	99.1	72.7	91.3	99.1	100.0	58.5	79.2	91.6	95.8	97.7	99.4	100.0
		n	437															
	All	Line	218	306	415	622	830	212	356	612	2,417	163	235	359	428	506	743	1,063
		Rate	49.6	65.4	78.2	91.5	95.0	60.2	82.6	94.5	100.0	48.2	67.2	82.7	87.8	91.8	96.3	98.4
		n	493															

Source: 2017 EDAM.

Poverty rates are percentages.

National poverty lines are DJF per adult equivalent, per day.

International 2011 PPP lines and percentile-based lines are DJF per-person, per-day.

All poverty lines are DJF in prices in Djibouti as a whole on average in May 2017.

Figure 10: (Obock, Arta): Poverty lines and head-count poverty rates by urban/rural/all in 2017

	Region/ Area	Line or Rate	*n*	Poverty lines and poverty rates														
				National					Intl. 2011 PPP				Percentile-based lines					
				Food	Minimum	100%	150%	200%	$1.90	$3.20	$5.50	$21.70	10th	20th	40th	50th	60th	80th	90th
Obock	Urban	Line	109	218	306	415	622	830	212	356	612	2,417	163	235	359	428	506	743	1,063
		Rate		4.9	19.5	31.8	48.8	62.9	11.6	37.4	62.1	100.0	3.3	20.8	38.7	44.7	54.2	69.3	84.9
	Rural	Line	366	218	306	415	622	830	212	356	612	2,417	163	235	359	428	506	743	1,063
		Rate		31.0	52.3	69.1	83.4	89.9	46.6	72.6	87.9	99.9	28.6	53.2	73.2	79.7	84.2	93.0	96.9
	All	Line	475	218	306	415	622	830	212	356	612	2,417	163	235	359	428	506	743	1,063
		Rate		21.5	40.4	55.6	70.9	80.2	33.9	59.8	78.5	100.0	19.4	41.4	60.7	67.1	73.3	84.4	92.6
Arta	Urban	Line	108	218	306	415	622	830	212	356	612	2,417	163	235	359	428	506	743	1,063
		Rate		2.7	3.6	16.0	42.2	57.2	3.5	21.3	55.3	98.5	0.9	3.6	21.3	31.7	45.6	69.8	91.4
	Rural	Line	372	218	306	415	622	830	212	356	612	2,417	163	235	359	428	506	743	1,063
		Rate		25.4	41.7	64.7	82.7	92.6	36.5	73.0	90.7	99.6	24.2	43.1	73.2	80.0	85.7	95.8	98.3
	All	Line	480	218	306	415	622	830	212	356	612	2,417	163	235	359	428	506	743	1,063
		Rate		19.4	31.6	51.8	72.0	83.2	27.8	59.3	81.3	99.3	18.0	32.6	59.4	67.2	75.1	88.9	96.4

Source: 2017 EDAM.

Poverty rates are percentages.

National poverty lines are DJF per adult equivalent, per day.

International 2011 PPP lines and percentile-based lines are DJF per-person, per-day.

All poverty lines are DJF in prices in Djibouti as a whole on average in May 2017.

3.2 Number of poor people in a single time period

Fulfilling a pro-poor mission depends not only on the *poverty rate* of in-coming participants but also on the *number* of poor in-coming participants. After all, a smaller program whose few participants have a higher poverty rate may serve fewer poor people than a larger program whose many participants have a lower poverty rate.[15]

The first step in estimating the number of poor people in one period is to estimate the number of household members in the population of in-coming households. In our two-household example with simple random sampling, this is the equal-weighted average of the number of people in the sampled households:

$$\frac{8+5}{1+1} = \frac{13}{2} = 6.5 \text{ people.}$$

The second step is to estimate the total number of people in the population of in-coming households. The example program has 1,000 in-coming households in its first calendar-year, each with an estimated 6.5 members. The estimated number of people in the households of in-coming participants is then 1,000 · 6.5 = 6,500.

The third and final step is to multiply the estimated poverty rate (here, 88.9 percent, or 0.889) by the estimated number of people in in-coming households (here, 6,500). This gives 6,500 · 0.889 ≈ 5,777 poor people (**Figure 9**).

All else constant, the *number* of in-coming participants who are poor is more important than the *share* of in-coming participants who are poor. Both estimates are useful,[16] but increasing the share who are poor is only a means to the end of increasing the number who are poor.

In turn, increasing the number of in-coming participants who are poor is only a means to the end of increasing the net reduction in the number of on-going participants who are poor.

---

[15] **Navajas *et al***., 2000.

[16] **Schreiner** (2014) tells how to report and analyze scorecard estimates.

3.3 Net changes in poverty rates across two time periods for on-going participants

The estimated net change in a population's poverty rate is the difference between the two estimated poverty rates at follow-up versus baseline.

Two sampling approaches are possible for the follow-up round after baseline:

- *One sample scored twice*: Score the same sample that was scored at baseline
- *Two independent samples*: Score a new sample from the same population that was scored at baseline

Given the scorecard's assumptions, both approaches are unbiased, although with all else held constant, scoring one sample twice has smaller margins of error than does scoring two independent samples.

3.3.1 Annual net change in poverty rates with one sample scored twice

When the follow-up sample is made up of the same households as the baseline sample,[17] then the estimated annual net change in the poverty rate of the population of on-going participants is the average-household-size-weighted average of the change in each scored household's poverty likelihood, divided by the household-size-weighted average of the years between each household's interviews.[18]

Continuing the earlier example, suppose that the first household at follow-up has seven members (rather than eight as at baseline) and is scored a second time on August 13, 2024, which is 1,157 days (about 3.17 years) after its first interview on June 13, 2021. Its score is now 27 (rather than 22), so its poverty likelihood for 100% of the national line is 83.2 percent (**Figure 1**).

Suppose that the second household now has six members (rather than five as at baseline) and is scored a second time on May 15, 2024, which is 1,050 days (about 2.88 years) after its first interview on June 30, 2021. Its score is now 35 (rather than 32), so its poverty likelihood has decreased from 81.1 to 63.4 percent.

[17] Or when the follow-up sample is a random sample of the baseline sample.

[18] Estimates of change do not directly adjust for the estimation error in estimates in a single period because—given the scorecard's assumptions—this error washes out when comparing follow-up with baseline. Error due to divergence from assumptions is unknown, and there is no direct way to adjust for it.

With poverty likelihoods expressed as proportions between 0 and 1, the average-household-size-weighted average of the change in each scored household's poverty likelihood is −12.2 percentage points:

$$\frac{\left(\frac{8+7}{2}\right)\cdot(0.832-0.913)+\left(\frac{5+6}{2}\right)\cdot(0.634-0.811)}{\left(\frac{8+7}{2}\right)+\left(\frac{5+6}{2}\right)}\approx\frac{-0.608+-0.974}{13}\approx-0.122.$$

The estimated head-count poverty rate decreased (improved) by 12.2 *percentage points* (not by 12.2 *percent*) between baseline and follow-up.

For clarity—and because the time between interviews varies across scored households—this estimate should be annualized by dividing it by the average-household-size-weighted average of years between the two interviews:

$$\frac{\left(\frac{8+7}{2}\right)\cdot 3.17+\left(\frac{5+6}{2}\right)\cdot 2.88}{\left(\frac{8+7}{2}\right)+\left(\frac{5+6}{2}\right)}\approx\frac{25.78+23.04}{13}\approx 3.05\text{ years}.$$

The annual, non-compounded rate of net change is then the percentage-point change in the poverty rate, divided by the average years between interviews: −12.2 ÷ 3.05 ≈ −4.0 percentage points per year.[19] The negative change means that poverty decreased (improved).[20]

In practice, the calculations should be done in an **app** or a spreadsheet like **Figure 11**.

[19] *Percentage points* are distinct from *percentages* (or *percents*). On the one hand, if the baseline poverty rate is 50.0 percent, and if there is a 10.0-*percent* annual reduction in the poverty rate, then the poverty rate after one year is 0.50 · (1 − 0.10) = 0.450 = 45.0 percent, and the poverty rate after two years is 0.45 · (1 − 0.10) = 0.405 = 40.5 percent. On the other hand, if there is a 10.0-*percentage-point* annual reduction in poverty, then the rate after one year is 0.50 − 0.10 = 0.40 = 40.0 percent, and the rate after two years is 0.40 − 0.10 = 0.30 = 30.0 percent.

[20] Of course, such a large annual reduction in poverty is unrealistic, but this is just an example to show how the scorecard can be used to estimate change.

Figure 11: Spreadsheet calculation of estimated annual net change in the head-count poverty rate and in the annual net number of poor people who rose above a poverty line with one sample scored twice

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	ID of direct	Interview date		Years between	Number of household members			Member-years	Score		Poverty likelihood (%)		Estimated net change in
2	participant	Baseline	Follow-up	interviews	Baseline	Follow-up	Average	between interviews	Baseline	Follow-up	Baseline	Follow-up	number of poor household
3	1V0276FZ7	13-Jun-2021	13-Aug-2024	3.17 = (C3-B3)/365	8	7	7.50 = (E3+F3)/2	23.77 = D3*G3	22	27	91.3	83.2	–0.608 = G3*(L3-K3)/100
4	2W3120ZG8	30-Jun-2021	15-May-2024	2.88 = (C4-B4)/365	5	6	5.50 = (E4+F4)/2	15.82 = D4*G4	32	35	81.1	63.4	–0.974 = G4*(L4-K4)/100
5				Average:	6.5 = AVERAGE(E3:E4)	6.5 = AVERAGE(F3:F4)	Sum:	39.60 = SUM(H3:H4)					–1.581 = SUM(M3:M4)
6													
7						Estimated net change in head-count poverty rate (percentage points), follow-up versus baseline:							–12.2 = M5/(E5+F5)*100
8													
9								Household-size-weighted average years between interviews:					3.05 = H6/(E5+F5)
10													
11							Estimated annual net change in head-count poverty rate (percentage points):						–4.0 = M7/M9*100
12													
13										Participating households at baseline:			1,000
14										Participating households at follow-up:			700
15													
16								Estimated average number of on-going participating people:					5,525 = (E5*M13+F5*M14)/2
17													
18								Estimated annual net change in the number of poor people:					–221 = M16*M11/100
19	Rows of data are sorted by the ID of the direct participant.												

3.3.2 Annual net change in the number of poor people with one sample scored twice

For a pro-poor program, the bottom line is *not* the annual net change in the poverty rate. Rather, the bottom line is the annual net change in the number of poor participants.

To calculate this, the first step is to estimate the average number of household members in the population of on-going participants from baseline to follow-up, accounting for drop-out. In the example here, the population in 2021 of in-coming households in the calendar-year 2021 cohort was 1,000. By the end of the follow-up period of calendar-year 2024, 300 had dropped out, leaving 700 from the 2021 cohort. If drop-out took place at a constant pace and was unrelated to changes in poverty,[21] then an estimate of the average number of on-going participating people is the equal-weighted average of the number of participating people among

---

[21] This assumption rarely holds. On the one hand, the households that benefit most from the program—and thus those for whom participation is most likely to cause a faster-than-otherwise decrease in poverty—may also be the least-likely to drop out, leading to too-high estimates of the reduction in poverty due to participation. On the other hand, households whose poverty decreases may be more likely to drop out if the benefits of continued participation fall as poverty decreases, leading to too-low estimates of impact. Unfortunately, there is no general way to adjust scorecard estimates to account for drop out that is related to changes in poverty. As in all decision-making, managers must use their experience and judgment to detect deviations from assumptions and then to account for them as best they can. This is true even though scorecard estimates are based on data and math. "Hard numbers" may not represent reality as accurately as they may seem to, and only a manager's knowledge of context can detect and account for this. Managers should discount unreliable estimates when they have reasoned, explicit arguments to do so (**Schreiner**, 2016a). Of course, discretion also opens the door to abuse; faced with unexpectedly low estimates of poverty reduction, managers might quietly sweep them under the rug or blame them on a slow economy (even though they would not attribute high estimates of poverty reduction to a roaring economy). Ironically and sadly, such attempts to make a program look good by hiding or excusing undesired results destroys the results' value as feedback, harming the program's ability to fulfill its mission. If a program's funders fail to act like owners, then its employees—not its participants—often become its *de facto* beneficiaries (**Schreiner**, 1997).

households interviewed at baseline and follow-up. In a given round, the number of participating people is the average household size for that round's interviewed households (in the example, 6.5 at baseline and 6.5 at follow-up), multiplied by the number of participating households in the population in the given round (1,000 at baseline and 700 at follow-up), divided by the number of survey rounds (two). This is $\frac{6.5 \cdot 1{,}000 + 6.5 \cdot 700}{1+1} = 5{,}525$ people.

The second and last step is to multiply the estimated annual change in the poverty rate (here, about −4.0 percentage points, or −0.040) by the estimated average number of on-going participants (here, 5,525). This gives an estimate of the annual net change in the number of poor people by 100% of the national line of $-0.040 \cdot 5{,}525 \approx -221$ people.[22] This negative change is a decrease (improvement) in poverty; there are about 221 fewer poor people in participating households in this cohort each year.

3.3.3 Estimating a program's impact

Estimating *change* is not the same as an estimating a program's *impact*. It stands to reason that program participation is a real force that does cause some change (be it an increase or decrease) in the poverty of its participants. At the same time, it is equally logical to expect that a large share of any change in participants' poverty is caused by the many non-program forces that also affect participants. On its own, the scorecard is like a bathroom scale; it can tell whether you lost weight in the past year, but not how much of the loss is due to eating right and exercising versus removing your coat and shoes.

This point is often forgotten, confused, or ignored, so it bears repeating: the scorecard estimates change, but it does not—on its own—identify the causes of change. In particular, estimating the impact of program participation requires knowledge or assumptions about what would have happened to participants if they had not been participants. This must come from beyond the scorecard.

What is a program manager to do? After all, decision-making hinges on forecasts of the expected impacts of possible choices; a manager cannot pretend that merely estimating change is helpful without also inferring some cause-and-effect relationship. Yet there are diminishing returns to improving inferences of impact. At a minimum, a program should compare its estimated annual net change in the poverty rate of its on-going participants to third-party estimates for the country

[22] This is a net figure; some start above the line and end below it, and vice versa.

overall or for the area where the program works (such as those in **Figure 10**). A program can also look for signs that participants value (or expect to value) its services. Is the number of in-coming participants high or increasing? Is the drop-out rate low or decreasing? Are drop-outs mostly due to dissatisfaction or graduation? Is participation voluntary, without being a condition for some other linked benefit? Is the program the sole provider in its niche and area?

In short, decision-makers in pro-poor programs are called to do what good decision-makers must always do: weigh data and knowledge from a number of perspectives and sources—including scorecard estimates, but not *only* scorecard estimates—to inform reasoned guesses as to more or less what share of observed changes are due to program participation. Of course, the inevitable need for human wisdom/art may be disingenuously invoked as a cover for decision processes that do not take a program's pro-poor mission to heart. This is why the "scientific method"—that is, being transparent about inputs and reasoning so as to facilitate productive review and debate—makes sense even (or perhaps especially) for business decisions.[23]

3.3.4 Annual net change in poverty rates with two independent samples

Instead of interviewing the same sample of households at both baseline and follow-up, a program could draw a second, independent sample of households from the same population as that from which the baseline sample was drawn.[24] The head-count poverty rate for on-going participants in this new follow-up sample is estimated in the same way as for the baseline sample.

Continuing the example, suppose that a third household and a fourth household are sampled at follow-up. The third household is interviewed on March 3, 2024. It has four members, a score of 29, and a poverty likelihood by 100% of the national line of 81.1 percent (**Figure 1**).

The fourth household is interviewed on April 4, 2024. It has seven members, a score of 37, and a poverty likelihood of 53.4 percent.

[23] **Schreiner** (2016a) and **Schreiner** (2014).
[24] By chance, some households may end up in both samples.

At follow-up, the estimated head-count poverty rate is calculated in the same way as at baseline, that is, as the household-size-weighted average of the poverty likelihoods of the sampled households:

$$\frac{4 \cdot 0.811 + 7 \cdot 0.534}{4 + 7} \approx \frac{3.24 + 3.74}{11} \approx 0.635 = 63.5 \text{ percent.}$$

The estimated annual net change in the head-count poverty rate of on-going participants is then the difference between the poverty-rate estimates at follow-up (63.5 percent) versus at baseline (87.4 percent),[25] divided by the difference (in years) between the household-size-weighted average of follow-up interview dates (March 23, 2024) versus the household-size-weighted average of baseline interview dates (June 19, 2021). These two average dates differ by about 1,008 days or about 2.76 years.

The estimated annual net change in the head-count poverty rate is the difference between the poverty-rate estimates at follow-up versus baseline, divided by the difference in the average years between interviews in the two rounds. For 100% percent of the national line, this is (63.5 − 87.4) ÷ 2.76 ≈ −8.7 percentage points per year.

In practice, the calculations should be done in an **app** or a spreadsheet like **Figure 12**.

[25] With two independent samples, the estimation error in each of the two single-period estimates washes out, so it is not explicitly included in the calculation. Thus, the figure here is 87.4 percent, not 87.4 −(−1.5) = 88.9 percent.

Figure 12: Spreadsheet calculation of estimated annual net change in a head-count poverty rate and in the annual net number of poor people who rise above a poverty line with two independent samples

	A	B	C	D	E	F	G	H
1	**Survey**	**ID of direct participant**	**Interview date**	**Number of household members**	**Interview date x Number of household members**	**Score**	**Poverty likelihood (%)**	**Estimated number of poor household members**
2	Baseline	1V0276FZ7	13-Jun-2021	8	16-Aug-2871 = C2*D2	22	91.3	7.30 = D2*G2/100
3	Baseline	2W3120ZG8	30-Jun-2021	5	01-Jul-2507 = C2*D2	32	81.1	4.06 = D3*G3/100
4	Follow-up	3XA76T21L	3-Mar-2024	4	11-Sep-2396 = C2*D2	29	81.1	3.24 = D4*G4/100
5	Follow-up	4Y8Y3EQS9	4-Apr-2024	7	01-Nov-2769 = C2*D2	37	53.4	3.74 = D5*G5/100
6		**Sum baseline:**		13 = SUM(D2:D3)				11.36 = SUM(H2:H3)
7		**Sum follow-up:**		11 = SUM(D4:D5)				6.98 = SUM(H4:H5)
8		**Average baseline:**		6.5 = AVERAGE(D2:D3)	19-Jun-2021 = SUM(E2:E3)/D6			
9		**Average follow-up:**		5.5 = AVERAGE(D4:D5)	23-Mar-2024 = SUM(E4:E5)/D7			
10								
11					**Estimated baseline poverty rate (%):**			87.4 = H6/D6*100
12					**Estimated follow-up poverty rate (%):**			63.5 = H7/D7*100
13								
14				**Average years between follow-up and baseline interviews:**				2.76 = (E9-E8)/365
15								
16			**Estimated annual net change in head-count poverty rate (percentage points):**					–8.7 = (H12-H11)/H14
17								
18					**Participating households at baseline:**			1,000
19					**Participating households at follow-up:**			700
20								
21				**Estimated average number of on-going participating people:**				5,175 = (D8*H18+D9*H19)/2
22								
23				**Estimated annual net change in the number of poor people:**				–448 = H21*H16/100
24	Rows of data are sorted by Survey, then by Interview date, then by the ID of the participant of record.							

3.3.5 Annual net change in the number of poor people with two independent samples

For a pro-poor program, the bottom line is not the annual net change in the poverty rate but rather the annual net change in the number of poor participants.

To calculate this, the first step is to estimate the average number of household members in the population of on-going households from baseline to follow-up, accounting for drop-out. In the example here, the population of the baseline 2021 cohort in 2021 is 1,000 in-coming households. By the end of the 2024 follow-up period, 300 households dropped out, leaving 700 from the 2021 cohort. If drop-out took place at a constant pace and was unrelated with changes in poverty, then an estimate of the average number of on-going participating people is the equal-weighted average of the number of participating people among households interviewed at baseline and follow-up. In a given round, the number of participating people is the average household size for that round's interviewed households (in our example, 6.5 at baseline and 5.5 at follow-up), multiplied by the number of participating households in the population in the given round (1,000 at baseline and 700 at follow-up), and divided by two (the number of rounds). This is

$$\frac{6.5 \cdot 1{,}000 + 5.5 \cdot 700}{1+1} = 5{,}175 \text{ people.}$$

The second and last step is to multiply the estimated annual net change in the head-count poverty rate (here, −8.7 percentage points, or −0.087) by the estimated number of on-going participants (here, 5,175). For 100% of the national line, this gives an annual net change in the number of poor people of about $-0.087 \cdot 5{,}175 \approx -448$ people per year. This negative change is a (non-compounded) decrease in poverty; the number of poor people in participating households decreases (improves) by 448 each year.

Given the scorecard's assumptions, both approaches to estimating change over time—one sample scored twice, and two independent samples—are unbiased. In general, the two approaches give different estimates (as in this example) because they interview different households at different times. All else constant, scoring one sample twice has smaller margins of error, but there may be context-specific reasons (related to operational costs or non-sampling errors) to score two independent samples.

4. How to design scorecard surveys and samples

To design a scorecard survey and its sample, a program must decide:[26]

- Who will do interviews
- Where and how to do interviews
- How to record responses and scores
- How to calculate estimates and report/analyze them
- Which participating households to interview
- How many participating households to interview
- How frequently to interview households
- Whether to track a population across multiple time periods
- Whether to interview the same participants twice

Decisions should follow from the program's goals, the business issues to be informed, and the budget. The central goals of the design are to:

- Inform issues that matter to the program
- Make sure that the sample is representative of a well-defined population

4.1 Who will do interviews

The enumerators who interview participating households must be trained to follow the **Interview Guide**. Enumerators may be:

- Program employees
- Contractors

4.2 Where and how to do interviews

Interviews should be:

- In-person, and
- At the sampled household's residence, and
- With an enumerator trained to follow the **Interview Guide**

This is the only recommended way. It follows Djibouti's DISED in the 2017 EDAM, so it provides the most-accurate and most-consistent data (and thus the best estimates).

[26] **IRIS Center** (2007) and **Toohig** (2008) also discuss this topic, covering sampling, budgeting, training, logistics, interviewing, piloting, and recording data.

Of course, it is possible to do interviews in non-recommended ways such as:

- Without an enumerator (such as by respondents' filling out paper or web forms on their own or responding to questions sent via e-mail, texts, or robo-calls)
- Away from home (such as a program's service point or a local meeting place)
- Not in-person (such as with an enumerator by phone)

While non-recommended methods may reduce costs, they also affect responses[27] and thus reduce the accuracy of estimates. This is why interviewing by a trained enumerator at the residence is recommended.

In some contexts—such as when a program's service agents do not already visit participants at their residences anyway as part of their normal work—a program might be willing to trade accuracy for a lower-cost, non-recommended approach. The business wisdom of this choice depends on context-specific factors that each program must judge for itself. To judge carefully, a program that is considering a non-recommended method should do a small test to see how responses differ when compared with a trained enumerator at the residence. Furthermore, all reporting should discuss the possible consequences of the non-recommended method.

4.3 How to record responses and scores

Responses and scores may be recorded by enumerators on:

- Paper, and then keyed into a database or spreadsheet at an office
- **Mobile devices**, and then uploaded to a database[28]

4.4 How to calculate estimates and report and analyze them

Analysts can calculate estimates by plugging data into spreadsheets (following the examples in Section **3**) or with the spreadsheet-based **ProveIt!™-brand reporting app**. **Schreiner** (2014) describes how to report and analyze scorecard estimates.

---

[27] **Schreiner**, 2015.

[28] **Scorocs** can help set up a system to collect **data with mobile devices** or to transfer data from paper forms into a database at the office. Support is also available for calculating estimates and for reporting and analysis.

4.5 Which participating households to interview

Given a population relevant for a particular business decision, the participating households to be interviewed can be:

- All relevant participants (a census)
- A representative sample of relevant participants
- All relevant participants in a representative sample of relevant service points and/or in a representative sample of relevant service agents
- A representative sample of relevant participants in a representative sample of relevant service points and/or in a representative sample of relevant service agents

A census is rarely necessary or appropriate, except for very small programs. Nevertheless, it may be less costly to interview all in-coming households as a standard part of in-take rather than managing who gets scored and who does not.

4.6 How many participating households to interview

If not determined by other factors, the number of participating households to interview can be derived from sample-size formulas to achieve a desired confidence level for a desired margin of error (**Annex 6**).

The focus of sample design, however, should be less on having enough interviews to achieve some arbitrary level of statistical significance and more on having a representative sample from a well-defined population that is relevant for informing decisions that matter to the program.

In practice, non-sampling errors in implementation and in the definition of the population often matter at least as much as errors due to smaller samples. Programs are often concerned about sample size, but as there is no point in deriving the ideal sample size unless proportional effort goes to mitigating other sources of error and then accounting for margins of error in the analysis stage. Of course, larger samples produce more-reliable estimates. In practice, however, almost no one reports or considers margins of error (even though they should), and estimates based on at least 1,000 interviews will rarely raise eyebrows (**Annex 6**).

4.7 How frequently to do interviews

The frequency of scorecard surveys can be:

- As a once-off project (precluding estimating change)
- Every three years (or at any other fixed or variable time interval, allowing estimating change)
- Each time a service agent visits a participant at home (allowing estimating change)

4.8 Whether to track a population across periods

The scorecard can estimate changes in poverty across periods, but not all programs want to do this. Some programs want to assess poverty only for in-coming participants.

4.9 Whether to interview the same participants twice

If a scorecard is to be applied more than once in order to estimate changes in poverty, then it can be applied with:

- One sample of participants, all of whom are scored at both baseline and follow-up
- Two samples of participants from the same population, with the first sample scored at baseline and the second sample scored at follow-up.

All else constant, scoring one sample twice gives estimates with smaller margins of error. This approach may also be less costly at follow-up, given that the sampled households have already been tracked down at baseline. Also, the follow-up round could be based on a random sample of the households interviewed at baseline.

4.10 Example of survey design in Bangladesh

An example set of choices is illustrated by the microfinance arms of BRAC and ASA, two pro-poor titans in Bangladesh who each have about 7 million participating households and who made plans to apply the scorecard for Bangladesh[29] with a sample of about 25,000 participants each.

Their design is that all loan officers in a random sample of branches score all participants each time these loan officers visit a homestead (about once a year) as part of their standard due diligence prior to loan disbursement. The loan officers record responses on paper in the field before sending the forms to a central office to be entered into a database and converted to poverty likelihoods for further analysis.

[29] **Schreiner**, 2013.

5. How to use scores for targeting

When a program uses the scorecard for segmenting (*targeting*) participants for differentiated treatment based on poverty, people in households with scores at or below a cut-off are labeled *targeted* and given one type of treatment. People in households with scores above a cut-off are labeled *non-targeted* and given another type of treatment.[30]

Households that score at or below a given cut-off should be labeled as *targeted*,[31] not as *poor*.[32]

Targeting is successful to the extent to which poor people truly below a poverty line are targeted (*inclusion*) or non-poor people truly above a poverty line are not targeted (*exclusion*). Of course, no poverty-assessment tool is perfect, and targeting is unsuccessful to the extent to which poor people truly below a poverty line are not targeted (*undercoverage*) or non-poor people truly above a poverty line are targeted (*leakage*).

---

[30] *Targeting status* (having a score at or below a targeting cut-off) is not the same concept as *poverty status* (having consumption expenditure below a poverty line). Poverty status is a fact that is defined by whether consumption expenditure is below a poverty line as directly measured by a survey. In contrast, targeting status is a program's policy choice that depends on a cut-off and on an indirect estimate from a scorecard.

[31] Other labels can be meaningful as long as they describe the segment and do not confuse targeting status (having a score below a program-selected cut-off) with poverty status (having consumption expenditure below an externally-defined poverty line). Examples include: *Groups A, B, and C*; *People with scores of 29 or less, 30 to 69, or 70 or more*; and *People who qualify for reduced fees, or who do not qualify*.

[32] After all, it is very unlikely that all targeted households are poor (their consumption expenditure is below a given poverty line). In the context of the scorecard, the terms *poor* and *non-poor* have specific definitions that are based on consumption expenditure and a poverty line. Using these same terms for targeting status is incorrect and misleading.

Figure 13 below depicts these four possible targeting outcomes. Targeting accuracy varies by the cut-off score. A higher cut-off has better inclusion and better undercoverage (but worse exclusion and worse leakage). In contrast, a lower cut-off has worse inclusion and worse undercoverage (but better exclusion and better leakage).

Figure 13: Possible targeting outcomes

Statut observé de la pauvreté		Segment de ciblage: Ciblé	Segment de ciblage: Non ciblé
	Pauvre	**Inclusion** Pauvre judicieusement ciblé	**Défaut de couverture** Pauvre par erreur non ciblé
	Non pauvre	**Fuite** Non pauvre par erreur ciblé	**Exclusion** Non pauvre judicieusement non ciblé

Programs should weigh these trade-offs when setting a cut-off. A formal way to do this is to assign net benefits—based on a program's values and mission—to each of the four possible targeting outcomes and then to choose the cut-off that maximizes the sum of net benefits.[33]

The five tables below show the scorecard's targeting outcomes by poverty line and by score cut-off for people in Djibouti:

- **Figure 14:** Inclusion (% people who are poor and correctly targeted)
- **Figure 15:** Undercoverage (% people who are poor but mistakenly not targeted)
- **Figure 16:** Leakage (% people who are not poor but mistakenly targeted)
- **Figure 17:** Exclusion (% people who are not poor and correctly not targeted)
- **Figure 18:** Hit rate (% people correctly targeted, that is, inclusion plus exclusion)

For a given score cut-off, each of the five figures below also show the share of all people who are targeted.

[33] **Adams and Hand**, 2000; **Hoadley and Oliver**, 1998.

Figure 14: Inclusion (% people who are poor and correctly targeted)

Targeting cut-off	% all people who are targeted	Inclusion (%)															
		National					Intl. 2011 PPP				Percentile-based lines						
		Food	Minimum	100%	150%	200%	$1.90	$3.20	$5.50	$21.70	10th	20th	40th	50th	60th	80th	90th
<=24	5.9	4.0	4.7	5.6	5.8	5.8	4.7	5.7	5.8	5.9	4.0	5.0	5.7	5.8	5.8	5.8	5.9
<=28	12.1	7.0	9.9	11.6	11.8	11.9	9.2	11.7	11.9	12.1	6.1	10.0	11.7	11.9	11.9	12.0	12.0
<=32	17.5	9.4	13.4	15.8	17.2	17.2	12.6	16.2	17.2	17.5	7.8	13.7	16.2	16.9	17.2	17.3	17.4
<=35	22.3	10.4	15.9	19.6	21.7	22.0	14.5	20.3	22.0	22.3	8.6	15.8	20.4	21.3	21.7	22.1	22.3
<=38	29.0	11.1	17.7	23.4	27.6	28.3	15.7	24.6	28.2	29.0	8.9	17.1	24.7	26.9	27.8	28.4	28.9
<=40	34.6	11.6	18.7	25.6	32.2	33.5	16.4	26.9	33.4	34.6	9.4	18.0	26.9	30.3	32.8	33.9	34.4
<=42	40.0	11.8	19.4	27.6	36.9	38.7	16.8	29.3	38.6	40.0	9.6	18.5	29.5	33.7	37.4	39.2	39.7
<=44	45.4	11.8	19.9	29.7	40.9	43.7	16.9	31.7	43.4	45.4	9.6	18.8	31.8	36.9	41.6	44.4	45.1
<=46	50.3	11.8	20.5	31.2	44.6	47.8	17.1	33.6	47.5	50.3	9.6	18.9	33.7	40.1	45.2	48.9	49.8
<=48	56.5	11.9	20.9	33.1	49.1	53.3	17.3	35.7	52.7	56.5	9.6	19.5	35.9	43.3	49.7	54.7	56.0
<=50	61.1	12.0	21.3	34.4	52.1	57.0	17.3	37.4	56.2	61.1	9.7	19.9	37.6	45.5	52.3	59.1	60.6
<=52	66.7	12.0	21.4	35.0	54.4	61.5	17.4	38.2	60.2	66.7	9.7	20.0	38.4	46.9	54.9	64.1	65.8
<=54	72.0	12.0	21.4	35.3	56.7	65.2	17.4	38.6	63.4	72.0	9.7	20.0	38.8	48.3	56.8	68.1	70.7
<=56	76.4	12.0	21.5	35.4	57.8	67.8	17.4	38.7	65.9	76.4	9.7	20.1	38.9	49.0	58.4	71.5	74.8
<=59	82.2	12.0	21.5	35.5	58.9	70.2	17.4	38.9	68.1	82.0	9.7	20.1	39.0	49.6	59.4	74.8	79.8
<=62	87.1	12.0	21.5	35.5	59.3	71.6	17.4	38.9	69.4	86.7	9.7	20.1	39.1	49.8	60.1	77.0	83.6
<=65	90.8	12.0	21.5	35.5	59.5	72.7	17.4	38.9	70.5	90.4	9.7	20.1	39.1	49.8	60.1	78.6	86.3
<=69	94.4	12.0	21.5	35.5	59.5	73.2	17.4	38.9	70.5	93.8	9.7	20.1	39.1	49.8	60.2	79.2	88.1
<=74	97.6	12.0	21.5	35.5	59.5	73.4	17.4	38.9	70.6	96.5	9.7	20.1	39.1	49.8	60.2	79.7	89.2

Scorecard applied to the validation sample.

Figure 15: Undercoverage (% people who are poor but mistakenly not targeted)

Targeting cut-off	% all people who are targeted	Undercoverage (%)															
		National					Intl. 2011 PPP				Percentile-based lines						
		Food	Minimum	100%	150%	200%	$1.90	$3.20	$5.50	$21.70	10th	20th	40th	50th	60th	80th	90th
<=24	5.9	8.1	16.8	29.9	53.8	67.7	12.7	33.3	64.9	92.4	5.8	15.1	33.4	44.1	54.5	74.1	83.6
<=28	12.1	5.1	11.6	24.0	47.7	61.6	8.2	27.2	58.8	86.3	3.7	10.0	27.4	38.0	48.4	67.9	77.5
<=32	17.5	2.6	8.1	19.7	42.4	56.3	4.9	22.7	53.4	80.9	2.0	6.4	22.8	32.9	43.0	62.6	72.1
<=35	22.3	1.7	5.7	16.0	37.9	51.6	2.9	18.6	48.7	76.0	1.2	4.2	18.7	28.6	38.5	57.8	67.2
<=38	29.0	1.0	3.9	12.1	32.0	45.3	1.8	14.3	42.5	69.3	0.9	3.0	14.4	23.0	32.5	51.5	60.6
<=40	34.6	0.4	2.9	10.0	27.3	40.1	1.1	12.0	37.3	63.7	0.3	2.1	12.2	19.6	27.5	45.9	55.1
<=42	40.0	0.2	2.1	7.9	22.7	34.9	0.7	9.6	32.1	58.3	0.2	1.6	9.6	16.2	22.8	40.7	49.8
<=44	45.4	0.2	1.6	5.9	18.6	29.9	0.6	7.3	27.3	52.9	0.2	1.3	7.3	12.9	18.6	35.5	44.4
<=46	50.3	0.2	1.1	4.3	14.9	25.7	0.4	5.3	23.1	48.0	0.1	1.1	5.4	9.8	15.0	31.0	39.7
<=48	56.5	0.1	0.6	2.4	10.5	20.3	0.1	3.2	18.0	41.8	0.1	0.5	3.2	6.6	10.5	25.1	33.5
<=50	61.1	0.1	0.3	1.1	7.4	16.5	0.1	1.5	14.5	37.2	0.1	0.2	1.5	4.4	7.9	20.8	28.9
<=52	66.7	0.1	0.1	0.6	5.1	12.0	0.1	0.7	10.4	31.7	0.1	0.1	0.7	2.9	5.3	15.8	23.7
<=54	72.0	0.1	0.1	0.3	2.9	8.4	0.1	0.3	7.3	26.3	0.1	0.1	0.3	1.6	3.4	11.7	18.8
<=56	76.4	0.0	0.0	0.2	1.8	5.7	0.0	0.2	4.8	21.9	0.0	0.0	0.2	0.9	1.9	8.3	14.7
<=59	82.2	0.0	0.0	0.0	0.7	3.3	0.0	0.0	2.6	16.4	0.0	0.0	0.0	0.3	0.9	5.1	9.7
<=62	87.1	0.0	0.0	0.0	0.2	2.0	0.0	0.0	1.2	11.6	0.0	0.0	0.0	0.1	0.1	2.8	5.9
<=65	90.8	0.0	0.0	0.0	0.0	0.8	0.0	0.0	0.2	7.9	0.0	0.0	0.0	0.1	0.1	1.3	3.2
<=69	94.4	0.0	0.0	0.0	0.0	0.4	0.0	0.0	0.1	4.5	0.0	0.0	0.0	0.1	0.1	0.6	1.4
<=74	97.6	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.1	1.8	0.0	0.0	0.0	0.1	0.1	0.2	0.3

Scorecard applied to the validation sample.

Figure 16: Leakage (% people who are not poor but mistakenly targeted)

Targeting cut-off	% all people who are targeted	Leakage (%)															
		National					Intl. 2011 PPP				Percentile-based lines						
		Food	Minimum	100%	150%	200%	$1.90	$3.20	$5.50	$21.70	10th	20th	40th	50th	60th	80th	90th
<=24	5.9	1.9	1.2	0.3	0.1	0.1	1.2	0.2	0.1	0.0	2.0	0.9	0.2	0.1	0.1	0.1	0.0
<=28	12.1	5.1	2.2	0.5	0.2	0.1	2.8	0.4	0.2	0.0	6.0	2.0	0.3	0.2	0.2	0.1	0.0
<=32	17.5	8.0	4.0	1.6	0.3	0.2	4.9	1.2	0.2	0.0	9.7	3.7	1.2	0.6	0.2	0.2	0.0
<=35	22.3	12.0	6.5	2.8	0.7	0.4	7.8	2.0	0.4	0.0	13.7	6.5	2.0	1.2	0.7	0.2	0.0
<=38	29.0	17.9	11.4	5.6	1.4	0.8	13.4	4.4	0.8	0.0	20.0	11.9	4.4	2.2	1.3	0.6	0.1
<=40	34.6	22.9	15.9	9.0	2.4	1.1	18.2	7.7	1.2	0.0	25.0	16.6	7.7	4.1	1.8	0.6	0.2
<=42	40.0	28.2	20.6	12.3	3.1	1.3	23.2	10.7	1.4	0.0	30.3	21.5	10.5	6.1	2.5	0.8	0.2
<=44	45.4	33.5	25.4	15.7	4.4	1.7	28.5	13.7	2.0	0.0	35.7	26.6	13.5	8.3	3.7	1.0	0.3
<=46	50.3	38.5	29.8	19.1	5.7	2.5	33.2	16.7	2.8	0.0	40.6	31.3	16.6	10.1	5.0	1.4	0.5
<=48	56.5	44.6	35.6	23.4	7.4	3.2	39.2	20.8	3.8	0.0	46.8	37.0	20.6	13.1	6.7	1.8	0.5
<=50	61.1	49.2	39.9	26.7	9.0	4.1	43.8	23.7	5.0	0.0	51.4	41.2	23.5	15.6	8.8	2.0	0.5
<=52	66.7	54.7	45.3	31.7	12.2	5.2	49.3	28.4	6.4	0.0	56.9	46.7	28.3	19.6	11.7	2.6	0.9
<=54	72.0	60.0	50.6	36.6	15.3	6.8	54.6	33.4	8.6	0.0	62.2	52.0	33.2	23.6	15.1	3.8	1.3
<=56	76.4	64.4	54.9	41.0	18.6	8.6	59.0	37.7	10.5	0.0	66.6	56.3	37.5	27.4	18.0	4.9	1.6
<=59	82.2	70.1	60.7	46.7	23.3	12.0	64.7	43.3	14.1	0.2	72.4	62.1	43.1	32.6	22.8	7.4	2.4
<=62	87.1	75.1	65.6	51.6	27.8	15.5	69.7	48.2	17.7	0.4	77.4	67.0	48.1	37.3	27.0	10.1	3.5
<=65	90.8	78.8	69.3	55.3	31.3	18.1	73.4	51.9	20.4	0.4	81.1	70.8	51.8	41.0	30.7	12.2	4.6
<=69	94.4	82.3	72.8	58.8	34.8	21.2	76.9	55.5	23.8	0.5	84.6	74.3	55.3	44.6	34.2	15.1	6.2
<=74	97.6	85.6	76.1	62.1	38.1	24.2	80.2	58.7	27.0	1.1	87.9	77.5	58.5	47.8	37.4	17.9	8.4

Scorecard applied to the validation sample.

Figure 17: Exclusion (% people who are not poor and correctly not targeted)

Targeting cut-off	% all people who are targeted	Exclusion (%)															
		National (def. 2017)					Intl. 2011 PPP (2017 def.)				Percentile-based lines (2017 def.)						
		Food	Minimum	100%	150%	200%	$1.90	$3.20	$5.50	$21.70	10th	20th	40th	50th	60th	80th	90th
<=24	5.9	86.0	77.3	64.1	40.3	26.3	81.4	60.8	29.2	1.7	88.3	79.0	60.7	50.0	39.6	20.0	10.5
<=28	12.1	82.9	76.3	64.0	40.2	26.3	79.7	60.7	29.2	1.7	84.3	77.9	60.6	49.9	39.6	20.0	10.5
<=32	17.5	79.9	74.4	62.8	40.2	26.2	77.7	59.9	29.1	1.7	80.6	76.2	59.7	49.6	39.5	20.0	10.5
<=35	22.3	76.0	72.0	61.6	39.8	26.1	74.7	59.1	29.0	1.7	76.6	73.5	58.9	49.0	39.1	19.9	10.5
<=38	29.0	70.0	67.1	58.8	39.0	25.7	69.2	56.7	28.5	1.7	70.2	68.0	56.6	47.9	38.5	19.5	10.3
<=40	34.6	65.0	62.5	55.4	38.1	25.4	64.3	53.4	28.1	1.7	65.2	63.3	53.2	46.0	37.9	19.5	10.3
<=42	40.0	59.8	57.9	52.1	37.3	25.2	59.4	50.4	27.9	1.7	60.0	58.5	50.4	44.0	37.3	19.4	10.3
<=44	45.4	54.4	53.1	48.8	36.0	24.8	54.1	47.4	27.4	1.7	54.6	53.4	47.4	41.9	36.0	19.2	10.2
<=46	50.3	49.5	48.6	45.3	34.8	24.0	49.3	44.3	26.6	1.7	49.6	48.6	44.4	40.1	34.7	18.7	10.0
<=48	56.5	43.4	42.9	41.1	33.0	23.2	43.4	40.3	25.5	1.7	43.5	43.0	40.3	37.1	33.0	18.4	10.0
<=50	61.1	38.8	38.6	37.7	31.4	22.4	38.8	37.4	24.4	1.7	38.9	38.7	37.4	34.6	31.0	18.1	10.0
<=52	66.7	33.3	33.2	32.8	28.2	21.3	33.3	32.7	22.9	1.7	33.3	33.2	32.7	30.5	28.0	17.5	9.6
<=54	72.0	28.0	27.9	27.8	25.2	19.6	28.0	27.7	20.7	1.7	28.0	28.0	27.7	26.5	24.6	16.3	9.2
<=56	76.4	23.6	23.6	23.4	21.8	17.9	23.6	23.4	18.8	1.7	23.6	23.6	23.4	22.8	21.7	15.2	8.9
<=59	82.2	17.8	17.8	17.8	17.1	14.5	17.8	17.8	15.2	1.5	17.8	17.8	17.8	17.6	16.9	12.7	8.1
<=62	87.1	12.9	12.9	12.9	12.7	10.9	12.9	12.9	11.7	1.3	12.9	12.9	12.9	12.8	12.8	10.0	7.0
<=65	90.8	9.2	9.2	9.1	9.1	8.4	9.2	9.1	9.0	1.3	9.2	9.2	9.1	9.1	9.1	7.9	5.9
<=69	94.4	5.6	5.6	5.6	5.6	5.2	5.6	5.6	5.5	1.1	5.7	5.6	5.6	5.6	5.5	5.0	4.2
<=74	97.6	2.4	2.4	2.4	2.4	2.2	2.4	2.4	2.3	0.6	2.4	2.4	2.4	2.3	2.3	2.2	2.1
<=100	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Scorecard applied to the validation sample.

Figure 18: Hit rate (% people correctly targeted, that is, inclusion plus exclusion)

Targeting cut-off	% all people who are targeted	Hit rate (= Inclusion + Exclusion) (%)															
		National					Intl. 2011 PPP				Percentile-based lines						
		Food	Minimum	100%	150%	200%	$1.90	$3.20	$5.50	$21.70	10th	20th	40th	50th	60th	80th	90th
<=24	5.9	90.0	82.0	69.8	46.1	32.2	86.1	66.5	35.0	7.6	92.3	84.0	66.4	55.8	45.4	25.9	16.4
<=28	12.1	89.8	86.2	75.6	52.1	38.2	88.9	72.4	41.1	13.7	90.3	87.9	72.3	61.8	51.5	32.0	22.5
<=32	17.5	89.3	87.9	78.7	57.3	43.5	90.3	76.1	46.3	19.1	88.3	89.9	75.9	66.5	56.7	37.2	27.9
<=35	22.3	86.3	87.8	81.2	61.5	48.1	89.3	79.4	50.9	24.0	85.2	89.3	79.3	70.2	60.8	42.0	32.8
<=38	29.0	81.1	84.8	82.3	66.6	54.0	84.9	81.4	56.7	30.7	79.1	85.1	81.2	74.8	66.2	47.9	39.2
<=40	34.6	76.7	81.2	81.0	70.3	58.8	80.7	80.3	61.5	36.3	74.7	81.3	80.2	76.3	70.7	53.4	44.8
<=42	40.0	71.6	77.3	79.7	74.2	63.8	76.2	79.7	66.5	41.7	69.5	77.0	79.9	77.7	74.7	58.6	50.0
<=44	45.4	66.2	73.0	78.5	77.0	68.4	70.9	79.0	70.8	47.1	64.2	72.1	79.2	78.8	77.7	63.6	55.4
<=46	50.3	61.3	69.1	76.6	79.4	71.8	66.4	77.9	74.1	52.0	59.2	67.5	78.1	80.2	80.0	67.6	59.8
<=48	56.5	55.3	63.8	74.2	82.1	76.5	60.7	76.1	78.2	58.2	53.1	62.5	76.3	80.3	82.7	73.1	66.0
<=50	61.1	50.7	59.9	72.2	83.5	79.4	56.1	74.8	80.6	62.8	48.5	58.6	75.0	80.0	83.3	77.2	70.6
<=52	66.7	45.2	54.6	67.7	82.6	82.8	50.6	70.9	83.1	68.3	43.0	53.2	71.1	77.4	83.0	81.6	75.4
<=54	72.0	39.9	49.3	63.1	81.9	84.8	45.3	66.3	84.1	73.7	37.7	47.9	66.5	74.8	81.4	84.4	79.9
<=56	76.4	35.6	45.0	58.9	79.6	85.7	41.0	62.1	84.7	78.0	33.4	43.6	62.3	71.8	80.1	86.8	83.6
<=59	82.2	29.9	39.3	53.3	76.0	84.7	35.3	56.6	83.3	83.4	27.6	37.9	56.8	67.2	76.3	87.6	87.8
<=62	87.1	24.9	34.4	48.4	72.0	82.5	30.3	51.7	81.1	88.0	22.6	33.0	51.9	62.6	72.9	87.1	90.6
<=65	90.8	21.2	30.7	44.7	68.7	81.1	26.6	48.0	79.4	91.7	18.9	29.2	48.2	58.9	69.2	86.5	92.2
<=69	94.4	17.7	27.2	41.2	65.1	78.4	23.1	44.5	76.0	95.0	15.4	25.7	44.7	55.4	65.7	84.2	92.4
<=74	97.6	14.4	23.9	37.9	61.9	75.6	19.8	41.3	72.9	97.1	12.1	22.5	41.4	52.1	62.5	81.9	91.4

Scorecard applied to the validation sample.

For an example cut-off of 38 or less in the previous figures, 29.0 percent of all people are targeted, and outcomes for 100% of the national line in the validation sample are:

- Inclusion: 23.4 percent are below the line and correctly targeted
- Undercoverage: 12.1 percent are below the line and mistakenly not targeted
- Leakage: 5.6 percent are above the line and mistakenly targeted
- Exclusion: 58.8 percent are above the line and correctly not targeted

Increasing the cut-off to 40 or less increases the share of of all people targeted to 34.6 percent. The higher cut-off improves inclusion and undercoverage but worsens leakage and exclusion:

- Inclusion: 25.6 percent are below the line and correctly targeted
- Undercoverage: 10.0 percent are below the line and mistakenly not targeted
- Leakage: 9.0 percent are above the line and mistakenly targeted
- Exclusion: 55.4 percent are above the line and correctly not targeted

Which cut-off is preferred depends on the sum of net benefits. If each targeting outcome has a per-person benefit or cost, then total net benefit for a given cut-off is:

Benefit per person correctly included	x	People correctly included	−
Cost per person mistakenly not covered	x	People mistakenly not covered	−
Cost per person mistakenly leaked	x	People mistakenly leaked	+
Benefit per person correctly excluded	x	People correctly excluded.	

To set an optimal cut-off, a program would:

- Assign benefits and costs to possible outcomes, based on its values and mission
- Tally total net benefits for each cut-off using **Figure 14** to **Figure 17** above for a chosen poverty line
- Select the cut-off with the highest total net benefit

The most difficult step is assigning benefits and costs to targeting outcomes. A pro-poor program that uses targeting—with or without the scorecard—should thoughtfully consider how it values successful inclusion and exclusion versus errors of undercoverage and leakage. It is healthy to go through a process of thinking explicitly and intentionally about how targeting outcomes are valued.

A common choice of benefits and costs is the *hit rate*, where total net benefit is the number of people correctly included or correctly excluded:

Hit rate =	1	x	People correctly included	−
	0	x	People mistakenly undercovered	−
	0	x	People mistakenly leaked	+
	1	x	People correctly excluded.	

Figure 18 shows the scorecard's hit rate for all cut-offs and poverty lines. For the example of 100% of the national line in the validation sample, total net benefit under the hit rate for a cut-off of 38 or less is 82.3 percent. That is, about four in five Djiboutians are correctly classified.

The hit rate weighs the successful inclusion of people below a poverty line the same as the successful exclusion of people above the line. If a program values inclusion more (say, twice as much) than exclusion, then it can reflect this by setting the benefit for inclusion to 2 and the benefit for exclusion to 1. Then the chosen cut-off will maximize (2 x people correctly included) + (1 x people correctly excluded).

As an alternative to assigning benefits and costs to targeting outcomes and then setting a score cut-off to maximize net benefits, a pro-poor program could set cut-offs based on aspects of targeting accuracy from the three figures below:

- **Figure 19:** Share of targeted people who are poor
- **Figure 20:** Poor people correctly targeted per non-poor person mistakenly targeted
- **Figure 21:** Share of poor people who are targeted

Figure 19: Share of targeted people who are poor

Targeting cut-off	% all people who are targeted	% targeted people who are poor															
		National					Intl. 2011 PPP				Percentile-based lines						
		Food	Minimum	100%	150%	200%	$1.90	$3.20	$5.50	$21.70	10th	20th	40th	50th	60th	80th	90th
<=24	5.9	67.6	80.0	95.2	97.8	98.3	80.2	95.8	98.1	100.0	66.8	84.1	96.2	97.8	97.8	98.5	100.0
<=28	12.1	57.8	82.0	96.3	98.2	98.8	76.4	96.9	98.7	100.0	50.3	83.1	97.1	98.3	98.5	99.2	99.9
<=32	17.5	53.9	76.9	90.8	98.4	98.8	72.0	92.9	98.7	100.0	44.5	78.6	93.0	96.8	98.6	99.0	99.9
<=35	22.3	46.4	71.0	87.5	97.1	98.4	65.0	90.9	98.2	100.0	38.5	71.0	91.1	94.8	97.0	99.0	99.8
<=38	29.0	38.2	60.8	80.7	95.1	97.4	54.0	84.9	97.1	100.0	30.7	58.9	85.0	92.4	95.5	97.8	99.5
<=40	34.6	33.7	53.9	74.0	93.2	96.8	47.3	77.7	96.5	100.0	27.4	52.0	77.8	88.0	94.7	98.1	99.6
<=42	40.0	29.5	48.5	69.1	92.2	96.8	42.0	73.3	96.5	100.0	24.0	46.3	73.7	84.6	93.7	98.1	99.4
<=44	45.4	26.0	44.0	65.5	90.3	96.3	37.2	69.8	95.7	100.0	21.2	41.4	70.1	81.7	91.8	97.9	99.4
<=46	50.3	23.5	40.7	62.1	88.7	95.1	33.9	66.7	94.5	100.0	19.1	37.7	67.1	79.9	90.0	97.2	99.0
<=48	56.5	21.1	37.0	58.6	86.9	94.3	30.6	63.3	93.2	100.0	17.1	34.6	63.6	76.8	88.1	96.9	99.1
<=50	61.1	19.5	34.8	56.3	85.2	93.3	28.4	61.2	91.9	100.0	15.8	32.5	61.5	74.5	85.7	96.7	99.2
<=52	66.7	18.0	32.1	52.5	81.6	92.3	26.0	57.3	90.3	100.0	14.5	29.9	57.6	70.5	82.4	96.1	98.7
<=54	72.0	16.6	29.8	49.1	78.8	90.5	24.1	53.6	88.0	100.0	13.4	27.8	53.9	67.1	79.0	94.7	98.2
<=56	76.4	15.8	28.1	46.4	75.6	88.8	22.8	50.6	86.2	100.0	12.7	26.2	50.9	64.2	76.4	93.6	97.9
<=59	82.2	14.7	26.2	43.2	71.6	85.5	21.2	47.3	82.8	99.7	11.8	24.4	47.5	60.4	72.3	91.0	97.0
<=62	87.1	13.8	24.7	40.8	68.1	82.2	20.0	44.6	79.7	99.5	11.2	23.0	44.8	57.2	69.0	88.4	96.0
<=65	90.8	13.3	23.7	39.1	65.5	80.1	19.2	42.8	77.6	99.5	10.7	22.1	43.0	54.8	66.2	86.5	95.0
<=69	94.4	12.8	22.8	37.7	63.1	77.5	18.5	41.2	74.7	99.4	10.3	21.3	41.4	52.8	63.8	84.0	93.4
<=74	97.6	12.3	22.1	36.4	61.0	75.2	17.9	39.8	72.3	98.9	10.0	20.6	40.0	51.0	61.7	81.6	91.4

Scorecard applied to the validation sample.

Figure 20: Poor people correctly targeted per non-poor person mistakenly targeted

Targeting cut-off	% all people who are targeted	Poor people targeted per non-poor person targeted															
		National					Intl. 2011 PPP				Percentile-based lines						
		Food	Minimum	100%	150%	200%	$1.90	$3.20	$5.50	$21.70	10th	20th	40th	50th	60th	80th	90th
<=24	5.9	2.1:1	4.0:1	19.9:1	44.1:1	56.2:1	4.0:1	22.7:1	50.6:1	All poor	2.0:1	5.3:1	25.2:1	44.1:1	44.1:1	67.0:1	All poor
<=28	12.1	1.4:1	4.5:1	25.7:1	56.1:1	80.8:1	3.2:1	31.4:1	75.0:1	All poor	1.0:1	4.9:1	33.7:1	58.8:1	67.8:1	122.2:1	1,093.8:1
<=32	17.5	1.2:1	3.3:1	9.8:1	60.9:1	79.0:1	2.6:1	13.1:1	75.1:1	All poor	0.8:1	3.7:1	13.4:1	29.9:1	70.0:1	102.5:1	1,584.0:1
<=35	22.3	0.9:1	2.4:1	7.0:1	33.3:1	60.7:1	1.9:1	10.0:1	55.7:1	All poor	0.6:1	2.4:1	10.2:1	18.3:1	32.6:1	97.3:1	557.7:1
<=38	29.0	0.6:1	1.6:1	4.2:1	19.2:1	37.1:1	1.2:1	5.6:1	33.5:1	All poor	0.4:1	1.4:1	5.7:1	12.1:1	21.5:1	45.3:1	194.7:1
<=40	34.6	0.5:1	1.2:1	2.8:1	13.7:1	30.5:1	0.9:1	3.5:1	27.6:1	All poor	0.4:1	1.1:1	3.5:1	7.3:1	17.7:1	52.3:1	222.3:1
<=42	40.0	0.4:1	0.9:1	2.2:1	11.8:1	30.0:1	0.7:1	2.7:1	27.4:1	All poor	0.3:1	0.9:1	2.8:1	5.5:1	15.0:1	52.2:1	160.6:1
<=44	45.4	0.4:1	0.8:1	1.9:1	9.3:1	25.8:1	0.6:1	2.3:1	22.1:1	All poor	0.3:1	0.7:1	2.3:1	4.4:1	11.3:1	46.4:1	180.4:1
<=46	50.3	0.3:1	0.7:1	1.6:1	7.8:1	19.3:1	0.5:1	2.0:1	17.2:1	All poor	0.2:1	0.6:1	2.0:1	4.0:1	9.0:1	34.3:1	99.7:1
<=48	56.5	0.3:1	0.6:1	1.4:1	6.6:1	16.5:1	0.4:1	1.7:1	13.8:1	All poor	0.2:1	0.5:1	1.7:1	3.3:1	7.4:1	31.1:1	109.7:1
<=50	61.1	0.2:1	0.5:1	1.3:1	5.8:1	13.9:1	0.4:1	1.6:1	11.3:1	All poor	0.2:1	0.5:1	1.6:1	2.9:1	6.0:1	29.1:1	117.4:1
<=52	66.7	0.2:1	0.5:1	1.1:1	4.4:1	11.9:1	0.4:1	1.3:1	9.3:1	All poor	0.2:1	0.4:1	1.4:1	2.4:1	4.7:1	24.8:1	74.0:1
<=54	72.0	0.2:1	0.4:1	1.0:1	3.7:1	9.6:1	0.3:1	1.2:1	7.4:1	All poor	0.2:1	0.4:1	1.2:1	2.0:1	3.8:1	17.8:1	55.8:1
<=56	76.4	0.2:1	0.4:1	0.9:1	3.1:1	7.9:1	0.3:1	1.0:1	6.3:1	2,781.3:1	0.1:1	0.4:1	1.0:1	1.8:1	3.2:1	14.7:1	45.6:1
<=59	82.2	0.2:1	0.4:1	0.8:1	2.5:1	5.9:1	0.3:1	0.9:1	4.8:1	354.4:1	0.1:1	0.3:1	0.9:1	1.5:1	2.6:1	10.1:1	32.8:1
<=62	87.1	0.2:1	0.3:1	0.7:1	2.1:1	4.6:1	0.3:1	0.8:1	3.9:1	207.6:1	0.1:1	0.3:1	0.8:1	1.3:1	2.2:1	7.6:1	23.9:1
<=65	90.8	0.2:1	0.3:1	0.6:1	1.9:1	4.0:1	0.2:1	0.7:1	3.5:1	216.5:1	0.1:1	0.3:1	0.8:1	1.2:1	2.0:1	6.4:1	18.9:1
<=69	94.4	0.1:1	0.3:1	0.6:1	1.7:1	3.5:1	0.2:1	0.7:1	3.0:1	171.8:1	0.1:1	0.3:1	0.7:1	1.1:1	1.8:1	5.2:1	14.1:1
<=74	97.6	0.1:1	0.3:1	0.6:1	1.6:1	3.0:1	0.2:1	0.7:1	2.6:1	89.8:1	0.1:1	0.3:1	0.7:1	1.0:1	1.6:1	4.4:1	10.7:1

Scorecard applied to the validation sample. "All poor" means "Only poor targeted".

Figure 21: Share of poor people who are targeted

Targeting cut-off	% all people who are targeted	% poor people who are targeted															
		National					Intl. 2011 PPP				Percentile-based lines						
		Food	Minimum	100%	150%	200%	$1.90	$3.20	$5.50	$21.70	10th	20th	40th	50th	60th	80th	90th
<=24	5.9	33.1	22.0	15.8	9.7	7.9	27.2	14.5	8.2	6.0	40.7	24.8	14.5	11.6	9.6	7.3	6.6
<=28	12.1	57.9	45.9	32.6	19.9	16.2	52.8	30.0	16.8	12.3	62.4	50.0	30.0	23.8	19.7	15.0	13.5
<=32	17.5	78.1	62.3	44.5	28.8	23.4	72.1	41.7	24.4	17.8	80.0	68.2	41.6	34.0	28.6	21.6	19.5
<=35	22.3	86.0	73.7	55.0	36.4	29.9	83.3	52.2	31.1	22.7	87.9	78.9	52.1	42.6	36.0	27.7	24.9
<=38	29.0	92.1	82.0	65.9	46.3	38.4	89.8	63.3	39.9	29.5	91.2	85.1	63.1	53.9	46.1	35.6	32.3
<=40	34.6	96.7	86.7	71.9	54.1	45.5	93.8	69.1	47.3	35.2	97.0	89.6	68.9	60.8	54.4	42.5	38.5
<=42	40.0	98.0	90.1	77.7	61.9	52.6	96.2	75.3	54.6	40.7	98.3	92.2	75.4	67.6	62.1	49.1	44.4
<=44	45.4	98.0	92.7	83.5	68.8	59.4	96.7	81.4	61.4	46.1	98.4	93.5	81.4	74.0	69.1	55.6	50.4
<=46	50.3	98.3	95.0	87.8	74.9	65.0	97.8	86.3	67.3	51.2	98.6	94.4	86.3	80.4	75.1	61.2	55.6
<=48	56.5	99.0	97.0	93.1	82.4	72.4	99.2	91.8	74.6	57.5	98.9	97.3	91.9	86.8	82.5	68.5	62.5
<=50	61.1	99.2	98.8	96.8	87.5	77.6	99.4	96.2	79.5	62.2	99.2	99.1	96.2	91.2	86.9	74.0	67.7
<=52	66.7	99.4	99.4	98.4	91.4	83.6	99.5	98.2	85.2	67.8	99.3	99.5	98.2	94.1	91.1	80.2	73.5
<=54	72.0	99.4	99.4	99.3	95.2	88.6	99.5	99.2	89.7	73.2	99.3	99.5	99.2	96.8	94.3	85.3	79.0
<=56	76.4	100.0	99.8	99.6	97.0	92.2	100.0	99.5	93.3	77.7	100.0	99.9	99.5	98.3	96.9	89.6	83.5
<=59	82.2	100.0	100.0	99.9	98.8	95.5	100.0	99.9	96.3	83.4	100.0	99.9	99.9	99.5	98.5	93.7	89.1
<=62	87.1	100.0	100.0	99.9	99.6	97.3	100.0	99.9	98.3	88.2	100.0	100.0	99.9	99.8	99.8	96.4	93.4
<=65	90.8	100.0	100.0	99.9	99.9	98.9	100.0	99.9	99.7	92.0	100.0	100.0	99.9	99.8	99.8	98.4	96.4
<=69	94.4	100.0	100.0	99.9	100.0	99.5	100.0	99.9	99.8	95.4	100.0	100.0	99.9	99.9	99.8	99.2	98.4
<=74	97.6	100.0	100.0	99.9	100.0	99.8	100.0	99.9	99.9	98.2	100.0	100.0	99.9	99.9	99.9	99.8	99.7

Scorecard applied to the validation sample.

For example, a pro-poor program could set a score cut-off to achieve a desired poverty rate—say, 70 percent—among targeted people. For 100% of the national line, targeting those who score 42 or less would target 40 percent of Djiboutians and give a head-count poverty rate among those targeted of 69.1 percent (**Figure 19**).

Figure 20 is a different way of looking at this same aspect of targeting accuracy. It shows the number of poor people correctly targeted (included) for each non-poor person mistakenly targeted (leakage). For 100% of the national line and a score cut-off of 42 or less, about 2.2 poor people are successfully targeted for every one non-poor person mistakenly targeted.

Alternatively, a pro-poor program might seek to target a desired share—such as half—of poor Djiboutians. **Figure 21** shows that a score cut-off of 35 or less would target 22.3 percent of all Djiboutians, a group in which 55.0 percent are poor by 100% of the national line.

Interview Guide

The excerpts quoted here are from:

DISED. (2018) "Manuel de l'Enquêteur: Quatrième Enquête Djiboutienne Auprès des Ménages pour les Indicateurs Sociaux (EDAM-IS)", [the *Manual*], **link**.

G1. Basic interview instructions

The scorecard can be filled out on paper in the field, with responses entered later in a spreadsheet or in your own database. Alternatively, Scorocs' cloud-based **data-collection tool** works in a web browser or as an app on Android phones, allowing data entry in the field or in the office. If there is no connection, then data is stored on the phone until it can be uploaded.

The scorecard should be administered by enumerators trained to follow this **Interview Guide**.

Fill out the scorecard header and the **Back-page Worksheet** first, following the directions found there.

In the scorecard header, fill in the number of household members in the space "Number of household members:" based on the list that you the enumerator made as part of the **Back-page Worksheet**.

Do not directly ask the first scorecard question ("In which region or arrondissement does the household live?"). Instead, fill in the response based on the knowledge that you the enumerator have of the region or arrondissement where the household lives.

In the same way, do not directly ask the second scorecard question ("How many members does the household have?"). Instead, mark the response based on the number of household members that you listed on the **Back-page Worksheet**.

Likewise, do not directly ask the third scorecard question ("In the last seven days, did any household member 5-years-old or older work at least one hour . . . ?"). Instead, mark the response based on what you the enumerator already know about the work status of the members of the household from when you compiled the **Back-page Worksheet**.

Ask the next two questions (“Can the male head (or the husband of the female head) read and write?” and “How many rooms are used by the household?”) directly of the respondent.

For the sixth scorecard question (“What is the main material of the floor?”), you the enumerator should try to observe the floor yourself, without asking the question directly of the respondent. If you can determine the main materal of the floor with complete certainty, then record the response and move on to the next question. Otherwise, go ahead and ask the question of the respondent.

Ask all of the four remaining questions directly of the respondent.

Read each question aloud word-for-word, in the order presented in the scorecard. Do not read the response options.

Study this **Interview Guide** carefully, and carry it with you while you work. Follow its instructions (including this one).

Remember that the respondent for the interview need not be the household member who is the participant of record with your program.

Likewise, the service agent to be recorded in the scorecard header is not necessarily the same as you the enumerator who does the interview. Rather, the service agent is the employee of the pro-poor program with whom the participant of record has an on-going relationship. If there is no such service agent or if you do not know if there is a service agent, then write “NONE” or “UNKNOWN” in those spaces in the scorecard header.

In general, do not leave blank spaces in the header. If the requested information does not exist, is not applicable, or is unknown, then write “NONE” or “UNKNOWN” in the blanks. This shows that you the enumerator tried to obtain the data. This may help avoid the need to return to the household later to try to get the data.

When you mark a response to a scorecard question, write the point value in the “Score” column and then circle the spelled-out response option, the pre-printed point value, and the hand-written points, like this:

5. How many rooms are used by the household?	A. One	0	
	B. Two	3	**3**
	C. Three or more	8	

When an issue comes up that is not addressed in this **Interview Guide**, its resolution should be left to the unaided judgment of you the enumerator and the respondent, as that apparently was the practice of Djibouti's DISED in the 2017 EDAM. That is, a program should not promulgate any definitions or rules (other than those in this **Interview Guide**) to be used by all its enumerators. Anything not explicitly addressed in this **Interview Guide** is to be left to the unaided judgment of each individual enumerator and the respondent.

Do not read the response options to the respondent. Instead, read the question, and then stop; wait for a response. If the respondent asks for clarification or otherwise hesitates or seems confused, then read the question again or provide additional assistance based on this **Interview Guide** or as you the enumerator deem appropriate.

In general, you should accept the responses given by the respondent. Nevertheless, if the respondent says something—or if you see or sense something—that suggests that the response may not be accurate, that the respondent is uncertain, or that the respondent desires assistance in figuring out how to respond, then you should read the question again and provide whatever help you deem appropriate based on this this **Interview Guide**.

While responses to questions in the scorecard are verifiable, in most cases you do not need to verify responses. You should verify only if something suggests to you that a response may be inaccurate and thus that verification might improve data quality. For example, you might choose to verify if the respondent hesitates, seems nervous, or otherwise gives signals that he/she may be lying, confused, or uncertain.

Likewise, verification may be called for if a child in the interviewed household or if a neighbor says something that does not square with a respondent's response. Verification may also be a good idea if you can see something yourself that suggests that a response may be inaccurate, such as a consumer durable that the respondent claims not to possess, or a child eating in the room or in the yard who has not been counted as a member of the household.

In general, the application of the scorecard should mimic as closely as possible the application of the 2017 EDAM by Djibouti's DISED. For example, interviews should done in-person by a trained enumerator at the residence of the participating household because that is what DISED did in the 2017 EDAM.

G2. Translation

You the enumerator should do the interview in a language which both you and the respondent speak and understand well.

The scorecard itself, the **Back-page Worksheet**, and this **Guide d'Entretien** are available in **French**, **Somali**, and **English**. There are not yet official, professional translations to other languages (such as Afar) spoken in Djibouti. Users should check **scorocs.com** to see what translations have been done since this writing. If there is not yet an official, professional translation to a desired language, then please contact **Scorocs** to arrange to collaborate on one.

G3. General interview guidance from the *Manual*

According to p. 3 of the *Manual*, "The quality, completeness, and accuracy of the data collected depends mainly on the quality of the work that you the enumerator do, as well as your responability and cooperativeness.

"It is your job as the enumerator to complete the [scorecard questionnaire] for all the sampled households that are assigned to you. Do not delegate to someone else any of your tasks, functions, or responsabilities."

G3.1 Do's

According to p. 3 of the *Manual*, you the enumerator should do as follows.

- "Do study this **Interview Guide** carefully so as to understand it completely
- Do follow the instructions and guidelines in this **Interview Guide**
- Do conduct interviews in-person directly with the household. Do the interviews yourself; your job as an enumerator cannot be delegated to anyone else
- Do be professional always. When interviewing, let your actions reflect the importance of your work
- Do dress appropriately and professionally when visiting interviewed households. A respectful appearance helps to motivate households to cooperate and to provide accurate responses"

G3.2 Don'ts

According to p. 4 of the *Manual*, you the enumerator do not do as follows.

- "Do not delegate your work as an enumerator to someone else
- Do not change any of the responses provided by the respondent
- Do not divulge, repeat, or discuss the information obtained from an interviewed household. Do not showed a filled-out scorecard questionnaire to anyone. Do not forget that a household's responses are to be kept strictly confidential
- Do not take anyone with you to an interview who has no business being there
- Do not pressure respondents. Likewise, do not encourage a respondent to respond in exchange for false promises or other inducements"

G3.3 Interview techniques

According to p. 7 of the *Manual*, "The instructions that follow focus on techniques for interviewing so as to ensure that the work is adequate and appropriate.

"For a successful project that obtains reliable, accurate data, it is crucial that you the enumerator understand the importance of knowing the scorecard questionnaire and this **Interview Guide** forwards and backwards.

"Keep in mind that the households that you the enumerator will interview will have varied reactions to being interviewed, according to their varied cultural backgrounds, attitudes, and norms for behavior. You will interact with households that may differ from those in your own experience in terms of their structure, socio-economic status, educational level, work status, customs, religion, and so on.

"Thus, you the enumerator must develop effective communication and understanding so as to establish healthy relationships with a variety of people. This will help you to succeed in the many and varied situations that you will face. In addition, you must build trust with interviewed households so that the scoring project will get the reliable, accurate data that it needs.

G3.4 Greeting the interviewed household

According to pages 7−8 of the *Manual*, "The first impression that you the enumerator make when you first meet the interviewed household is crucial for the success of the project. Your later work will be affected by your appearance, your attitude, and what your say right at that start of the relationship. Be sure to dress appropriately and professionally.

“Once you the enumerator locate the sampled household to be interviewed, ask to speak with the head (or his/her spouse). Greet the person in a kind, friendly manner, introduce yourself, and explain that you are there as part of a survey intended to help your program to learn more about its participants, for which you would appreciate the household’s cooperation.

“An example script that you could use to introduce yourself is as follows:

> Good morning, my name is <YOUR NAME>. I am working as an enumerator for <YOUR PROGRAM>. Here is my identification badge. I would like to speak with the head of the household, please.

“You the enumerator should maintain a friendly demeanor with the respondent while also radiating self-confidence. If you give the impression that you are nervous or unsure of yourself, then the respondent may be less willing to cooperate, pay attention, or make an effort to participate.

“You the enumerator should always do your best to maintain an even keel throughout the interview. If, for some reason, the respondent is tired or gets annoyed, then pause the interview for a few moments. Do not continue until the respondent has calmed down and is ready to begin again.”

G3.5 Effective communication

According to page 8 of the *Manual*, “After you the enumerator introduce yourself, explain that the survey involves a sample of households that are participants with your program and that this household’s cooperation is greatly appreciated. Emphasize that the data collected will be kept strictly confidential. This point is crucial to head off the fear that the data may be misused and that the household will later regret its participation. Stress that the household’s data will only be used for statistical purposes and that no one will be able to link the household with its responses.

“Keep in mind that data quality is lowest at the start of the interview; it is then that the respondent’s attention, communication, confidence, and engagement are weakest. Your task as the enumerator is to continually increase the respondent’s engagement and then to maintain it at the highest level until the last question is asked. Data quality is affected by your rhythm, tone, and speed when asking questions, as well as your knowledge of the questions and their sequencing. If you ask questions in a monotone or if you sound nervous, then there is greater risk that the respondent will lose interest and provide less-accurate data.

“Do not give the impression that you think you are a big shot just becaue you are an enumerator. Instead, be open, friendly, and yet decisive, so that you come off as someone with who knows what you are doing. Do not be bossy nor aggressive. Communication blossoms when the respondent sees you as honest and on top of your game.”

G3.6 During the interview itself

According to pages 8−9 of the *Manual*, “Once the interview starts, you the enumerator should always follow these instructions:

- “Try to get a response for each question
- Carefully and clearly record the responses
- Ask the questions in the scorecard questionnaire. You can also ask follow-up questions to cross-check responses, if you believe it might improve accuracy
- Be polite. You should knock on the door; do not enter until you are invited in, and do not sit down until you have been offered a seat. If your host forgets to invite you to sit down, then you should politely ask for permission
- Do not say anything unless you are certain it is correct. It is better to appear uninformed, yet honest. Avoid any topics or attitudes that could provoke a discussion or argument with the respondent. Stick to the topics in the scorecard questionnaire
- When trying to convince the household to agree to be interviewed, do not promise any compensation nor other incentives
- Never reveal surprise nor any other reaction after hearing a household’s response, whether by your tone of voice or the look on your face
- Stick strictly the order and wording of the questions as they appear in the scorecard questionnaire
- Scrupulously follow the scorecard questionnaire and this **Interview Guide**. Any divergence risks destroying the consistency (and thus the usefulness) of the data for a given question. Read the questions exactly as they are written in the scorecard questionnaire, without any changes. If the respondent does not understand the question the first time, then re-read it. If the respondent still does not understand, then carefully explain the purpose and intent of the question, being careful not to change its original meaning nor to suggest a response
- Read the questions without pressuring the respondent and without suggesting any particular response. For example, never say something like, “You worked last week, right?” Never assume that you know what a response will be; let the respondent tell you his/her response

- Keep in mind that an interview is made up of questions, responses, pauses, and moments of silence. When you read questions, keep a consistent rhythm, giving the respondent time to reflect and give an accurate, careful response. You should assess how well the respondent understands the questions, and adjust your reading speed accordingly. Pronounce each word clearly
- Try to do the interview out of earshot of third parties who are not members of the interviewed household; the presence of third parties can lead the respondent to give inaccurate responses. In practice, however, it is not always possible to isolate the household for the interview. In these cases, you the enumerator should tactfully ask the third parties to leave you alone with the interviewed household because the respondent him/herself and the other members of the household are supposed to be the ones who reply
- After asking the last question, carefully review the scorecard questionnaire to ensure that all questions have a response marked. If there is an error, then take advantage of your being with the interviewed household right then and there to make the correction
- Before taking your leave of the interviewed household, be sure to thank the respondent and the other household members profusely for their cooperation"

G3.7 Who should be the respondent?

According to p. 15 of the *Manual*, "The head of the household is the preferred respondent. If the head is unavailable, then the other members of the household should choose one member from amongst themselves to respond in the place of the head. The respondent should know the [age and work status] of all the members of the household. Members who are not the respondent can still provide supplemental information in the interest of accuracy, in particular for questions about themselves."

"Usually, the main respondent will be the head of the household or his/her spouse/conjugal partner. Nevertheless, in some cases, some other adult member of the household may serve as the respondent."

G3.8 Who is the head of the household?

Note that the head of the household may or may not be the household member who is the participant of record with your program.

Every household has one (and only one) head. The head of the household must be a member of the household. A person cannot be the head of more than one household.

According to p. 15 of the *Manual*, "The members of the household themselves will identify their head. It is the person who they name as the head when you the enumerator ask "Who is the head of the household?" Usually, the head is the member who is in charge of managing the household's money, who is the oldest, and whose authority as head is acknowledged by all the other members of the household. The head must be a member of the household. A woman can be the head of the household."

G4. Guidelines for each question in the scorecard

G4.1 In which region or arrondissement does the household live?

A. Tadjourah
B. Dikhil
C. Ali Sabieh
D. Djibouti-ville (2nd or 4th arrondissement)
E. Obock, or Djibouti-ville (3rd ou 5th arrondissement)
F. Arta, or Djibouti-ville (1st arrondissement)

Unless you have to, do not directly ask this question of the respondent. Instead, fill in the response based on your knowledge of the region and arrondissement where the household lives.

G4.2 How many members does the household have?

A. Eight or more
B. Seven
C. Six
D. Five
E. Four
F. Three
G. Two
H. One

Do not directly ask this question of the respondent. Instead, mark the response based on the number of household members that you listed on the **Back-page Worksheet**.

According to p. 12 of the *Manual*, "A household is a group of people (regardless of blood or marital relationships) who recognize one member as the head of their household. Members of the household share all or part of their income and share expenses. Usually, members of a household live together in a residence and eat at least one meal together a day. The residence occupied by members of a household may be part of one building, all of one building, or more than one buildings. One person who lives alone (unmarried, widowed, or divorced) independently with his/her own income is the head of a one-person household.

"This definition means that the membership of a given person in a given household requires that all of the following four criteria be met:

- "*Sharing all or part of income*: Expenses covered by income provided by one member of the household benefit all the members of the household to some degree, and
- *Recognition of the authority of the head of the household*. This is probably the most-objective of the four criteria, and
- *Living together in the same residence*, and
- Eating meals together. In urban areas, this usually means the evening meal

"A *household* is not necessarily the same thing as a *family* (in the sense of a nuclear family made up only of a father, a mother, and their own children). Still, most households are indeed nuclear families.

“If the interviewed household has domestic servants who share their income and their residence with the interviewed household, then those domestic servants are counted as members of the interviewed household.”

According to p. 15 of the *Manual*, one of two additional criteria must also be met for a given person to be a member of a given household:

- “The person must have lived with the household for more than six month as of the date of the interview, or
- The person must currently live with the household and intend to remain with the household for a total duration of at least six months. Common examples include new spouses who have joined the household, newborns, and adopted children

“If someone—for whatever reason—has lived apart from the household during the six months before the date of the interview, then that person is not counted as a member of the household.

“You the enumerator should make a list of all the members of the household (including those who happen to be temporarily absent on the day of the interview but who otherwise would qualify as members of the interviewed household) [using the **Back-page Worksheet**]. Make the list in this order:

- “Head of the household
- Spouse/conjugal partner of the head of the household
- Children of the head of the household and their spouses/conjugal partners
- Children with one or more parents who are not members of the interviewed household (for example, orphans, children of a father who no longer lives with the interviewed household, and so on)
- Other relatives by blood or by marriage, as well as people who do not have a blood nor marital relationship with the head of the household”

“Be sure to count babies, elderly people, and bed-ridden people in the household. Also remember to count people who are temporarily absent on the day of the interview but who otherwise would qualify as members of the interviewed household.”

Pages 25−26 of the *Manual* list a number of other rules for determining whether a given person is a member of a given household.

“Someone (such as a son or daughter, brother, sister, relative, or an unrelated person) has his/her own personal income (for example, from a wage/salary job, a small store, and so on) that is partly turned over to the head of the interviewed household in which the person sleeps and eats is counted as a member of the interviewed household if he/she turns over more than half of his/her income and allows the head of the interviewed household to use that money as the head sees fit. However, this person is not a member of the interviewed household if he/she pays only for room and board, keeping the rest of his/her income for his/her own use.

“Someone who is completely dependent on the household (for example, an adult, unmarried child who lives with his/her parents, or a young man who has moved to the city to look for a job and who lives with a relative) for all his/her basic needs is counted as member of the household where he/she lives.

“A parent or an aging relative with low or no income who lives with (and is supported by) another relative is a member of the household where he/she lives.

“A woman and her children who live and eat apart from their husband/father (regardless of whether he has other wives) who stops by from time to time (to visit, to share a meal, or to pick up food that the woman sometimes prepares for him) are counted as members of a household in which the woman is the head. The husband in not counted as a member of the wife’s household and thus cannot be the head of the wife’s household.

“A polygamous man who lives with some or all of his wives in a single compound is counted as a member (along with his co-resident wives) of the household where he lives.

“The wives of a polygamous man who live with their children in separate compounds are counted as separate households in which the wives are the heads. The polygamous man is counted as a member (and the head, if he qualifies as such) of the household where he stays.

“A household temporarily without any resources that is ‘rescued’ by another household (for example, a household in the same compound) from whom the devastated household receives money for food (or from whom the devastated household receives meals, shelter, and other basic needs as gifts) is counted as a separate household apart from its benefactor household.

“Young people living together as roommates (for example, young students or college students, or young people from a rural area who have moved to the city to seek work) are counted as:

- A single household (if they share both meals and the residence)
- A single household (if they share meals and if they live in separate residences in a single compound)
- Separate households (if they share a residence but do not share meals)

“A maid or other domestic servant is counted as:

- A member of his/her employer’s household (if the employer provides for his/her food, shelter, and other basic needs)
- A member of some household other than his/her employer’s household (if the employer provides food but not shelter)
- A member of some household other than his/her employer’s household (if the employer does not provide food, regardless of whether the employer provides shelter)”

G4.3 In the last 7 days, did any household member 5-years-old or older work at least one hour for a business or in self-employment, produce a good or service for another household or to make money for his/her own household, or work as a paid or unpaid intern?

A. No
B. Yes

Do not directly ask this question of the respondent. Instead, mark the response based on the work status of the members of the household based on the information that you the enumerator recorded as part of the **Back-page Worksheet**.

The *Manual* does not have any additional information about this question.

G4.4 Can the male head (or the husband of the female head) read and write?

A. No
B. No male head (and the female head has no husband in the household)
C. Yes

Note that the head of the household may or may not be the household member who is the participant of record with your program.

Every household has one (and only one) head. The head of the household must be a member of the household. A person cannot be the head of more than one household.

According to p. 15 of the *Manual*, "The members of the household themselves will identify the head of the household. It is the person who they name as the head when you the enumerator ask "Who is the head of the household?" Usually, the head is the member who is in charge of managing the household's money, who is the oldest, and whose authority as head is acknowledged by all the other members of the household. The head must be a member of the household. A woman can be the head of the household."

Remember that you the enumerator already know the name of the male head (or the husband of the female head) and whether he exists from compiling the **Back-page Worksheet**. Thus, if there is a male head (or a husband of the female head), then do not mechanically ask, "Can the male head (or the husband of the female head) read and write?" Instead, use the actual name of the male head (or the husband of the female head), for example: "Can Mohamed read and write?" If there is no male head (nor husband of the female head), then do not ask the question of the respondent at all. Instead, mark "B. No male head (and the female head has no husband in the household)" and go to the next question.

For the purposes of the scorecard, the *male head* (*or the husband of the female head*) is defined as:

- The household head, if the head is male
- The spouse/conjugal partner of the household head, if the head is female and if her spouse/conjugal partner is a member of the interviewed household
- Non-existent, if the head is female and if she does not have a spouse/conjugal partner who is a member of the interviewed household

The *Manual* does not have any additional information about this question.

G4.5 How many rooms are used by the household? (Exclude hallways, bathrooms, closets, storerooms, and rooms used only for business purposes)

A. One
B. Two
C. Three or more

According to p. 20 of the *Manual*, "Do not count rooms used for business (beauty salon, office, and so on)."

The *Manual* does not have any additional information about this question.

G4.6 What is the main material of the floor?

A. Dirt, or other
B. Cement
C. Wood, or tile

If you can observe the response to the sixth scorecard question about the main material of the floor, then record it without asking the respondent. Ask the respondent only if you cannot observe it yourself with complete certainty.

The *Manual* does not have any information about this question.

G4.7 What is the household's main source of water?

A. River, creek, rainwater, dam, reservoir, underground cistern, unimproved well, improved well (without pump), borehole (well with pump), water truck, or public standpipe, or other
B. Piped by ONEAD to the residence's yard, or piped from a borehole
C. Piped by ONEAD inside of the residence

The *Manual* does not have any information about this question.

G4.8 Does the household have a TV?

A. No
B. Yes

According to p. 21 of the *Manual*, “Count TVs in good working order that the interviewed household owns or uses.”

The *Manual* does not have any additional information about this question.

G4.9 Is the main source of lighting for the household electricity from EDD?

A. No
B. Yes

The *Manual* does not have any information about this question.

G4.10 How many cell phones or radio phones does the household have?

A. None
B. One
C. Two or more

According to p. 21 of the *Manual*, “Count cell phones or radio phones in good working order that the interviewed household owns or uses.”

The *Manual* does not have any additional information about this question.

Technical Annexes: Overview

The technical annexes cover advanced aspects of the scorecard. While program managers can skip the annexes and still benefit from using the scorecard, understanding the details will increase the usefulness of scorecard estimates and improve implementation and interpretation.

Annex 1 Data used for construction and validation

Djibouti's Direction de la Statistique et des Ètudes Démographiques (DISED) fielded the 2017 Enquête Djiboutienne Auprès des Ménages (EDAM) with 4,474 households in 2017 in two stages: from April 17 to June 4, and from November 8 to December 23. The 2017 EDAM is Djibouti's most-recent available national household expenditure survey.

Questions and points for the scorecard are selected (*constructed*) based on data from a random three-fifths of the 4,474 households in the 2017 EDAM. These same three-fifths of households are also used to associate (*calibrate*) scores with poverty likelihoods for all poverty lines.

Data from the other two-fifths of households from the 2017 EDAM is used to test (*validate*) the scorecard's accuracy for one-period estimates of poverty rates *out-of-sample*, that is, with data that is not used in construction nor calibration. Data from those same two-fifths of households are also used for out-of-sample validation of targeting accuracy.

Annex 2 Definition of *poverty*

A household's *poverty status* as poor or non-poor depends on whether its consumption expenditure (DJF per adult equivalent per day or per person per day) is below a given poverty line. Thus, a definition of *poverty* is a poverty line together with a measure of consumption expenditure from the 2017 EDAM.[34]

Because pro-poor programs in Djibouti may want to use different or various poverty lines, the scorecard supports 16 lines:

- Food (seuil alimentaire) line
- Minimum (seuil extrême) national line
- 100% of the (seuil global) national line
- 150% of the (seuil global) national line
- 200% of the (seuil global) national line
- $1.90/day 2011 PPP line
- $3.20/day 2011 PPP line
- $5.50/day 2011 PPP line
- $21.70/day 2011 PPP line
- 10th-percentile line
- 20th-percentile line
- 40th-percentile line
- 50th-percentile line
- 60th-percentile line
- 80th-percentile line
- 90th-percentile line

A2.1 National poverty lines

Djibouti has three official poverty lines: a food (*seuil alimentaire*) line, and two food+plus-non-food lines, one of which is a lower minimum (*seuil extrême*) national line, and the second of which is a higher (*seuil global*) national line.[35] They are derived according the cost-of-basic-needs approach.[36]

---

[34] **DISED** (2018, pp. 3−5) describes the measurement consumption expenditure in the EDAM, which follows **Deaton and Zaidi** (2002).

[35] **DISED** (2018)

[36] **Ravallion**, 1998.

A2.1.1 Food (*seuil alimentaire*) line

Djibouti's food poverty line (*seuil alimentaire*) is the cost of a minimum standard of 2,115 Calories per day for a basket of 44 food items that account for 81 percent of the total food consumption expenditure observed for the population of Djibouti in the 2017 EDAM (prices for Djibouti as a whole in May 2017).[37] Each item's share in the basket is its observed share of the total food consumption expenditure for the 44 included items. The resulting food line is about DJF218 per adult equivalent per day, giving a head-count poverty rate for all-Djibouti of 11.3 percent (**Figure 10**).

A2.1.2 Minimum (*seuil extrême*) national line

The minimum (*seuil extrême*) national poverty line is the food line, plus a minimum standard for non-food consumption expenditure, taken for this minimum line as the observed non-food expenditure for households in the 2017 EDAM whose *total* expenditure is about the same as the food line.[38] The minimum national line is about DJF306 per adult equivalent per day, with a head-count poverty rate for all-Djibouti of 21.1 percent (**Figure 10**).

A2.1.3 National (*seuil global*) line

The national (*seuil global*) poverty line is the food line, plus a less-stingy standard for non-food consumption expenditure: the observed non-food expenditure for households in the 2017 EDAM whose *food* expenditure is at the food line.[39] The national national line (usually referred to here as "100% of the national line") is about DJF415 per adult equivalent per day, with a head-count poverty rate for all-Djibouti of 35.8 percent (**Figure 10**).[40]

Both the food (*alimentaire*) and minimum (*extrême*) poverty lines are very stingy; people who are poor by these lines do not get enough to eat. Thus, it is recommended that pro-poor programs in Djibouti—if they assess poverty with an official DISED line—use 100% of the national (*global*) line.

150% of the national line and 200% of the national line are multiples of 100% of the (*global*) national line.

---

[37] **DISED** (2018, pp. 31−32)
[38] **DISED** (2018, pp. 32−33)
[39] **DISED** (2018, pp. 33−34)
[40] The all-Djibouti head-count poverty rates reported here for the minimum (*extrême*) line and the national (*global*) line match those in **DISED** (2018, p. 12), suggesting that this paper uses the same data and calculations as DISED did.

A2.2 International 2011 PPP poverty lines

The World Bank tracks world-wide poverty with four 2011 PPP poverty lines:[41]

- $1.90/day Low-income countries (the international "extreme poverty" line)
- $3.20/day Lower-middle-income countries
- $5.50/day Upper-middle-income countries
- $21.70/day High-income countries

PPP lines partly control for differences in purchasing power across countries due to the fact that non-tradable goods and services are usually less costly in poorer countries while tradables are more costly. PPP adjustments improve the comparability of poverty estimates across countries.

The World Bank classifies Djibouti as an lower-middle income country. The most relevant international line is probably $3.20/day, as about 40 percent of Djiboutians are below that line in 2017.[42]

International 2011 PPP lines for Djibouti are derived from:

- 2011 PPP (revised) exchange rate for Djibouti for "individual consumption expenditure by households":[43] DJF100.624 per $1.00
- Average all-Djibouti Consumer Price Index[44] (CPI):
  - Calendar-year 2011: 105.07
  - May 2017: 116.28

DISED does not adjust its poverty lines for price differences by region or arrondissement, but it does account for prices in the two periods of field work in the 2017 EDAM by adjusting prices to May 2017.

Given this, the $1.90/day 2011 PPP line for all households in Djibouti is:

$$\$1.90 \cdot 2011 \text{ PPP factor} \cdot \frac{CPI_{2017EDAM}}{CPI_{2011}} = \$1.90 \cdot 100.624 \cdot \frac{116.28}{105.07} = \text{DJF}211.58.$$

The corresponding head-count poverty rate is 17.0 percent (**Figure 10**).[45]

The 2011 PPP lines for $3.20/day, $5.50/day, and $21.70/day are multiples of the $1.90/day line.

---

[41] **Jolliffe and Prydz**, 2016; **Ferreira *et al.***, 2016.
[42] **World Bank**, 2021.
[43] **World Bank**, 2020, Table E.3, column 13, p. 135.
[44] Base = 100 in calendar-year 2010, **link**.
[45] This $1.90/day line and rate match those of **PovcalNet**.

A2.3 Percentile-based poverty lines

The scorecard supports percentile-based poverty lines.[46] This facilitates a number of types of analyses. For example, the second-quintile (40th-percentile) line might be used to help track Djibouti's progress toward the **World Bank**'s (2013) goal of "shared prosperity/inclusive economic growth", defined as income growth among the bottom 40 percent of the world's people.

Analyzed together, the four quintile lines (or all seven percentile lines) can also be used to look at the relationship of consumption expenditure with health outcomes (or anything else related with the distribution of consumption expenditure). The scorecard thus offers an alternative for health-equity analyses that typically have used an asset index (such as that supplied with the data from the Demographic and Health Surveys) to compare an estimate of socio-economic status with health outcomes.[47]

Of course, relative-wealth analyses are also possible with scores from the scorecard. But support for relative consumption expenditure lines allows for a more straightforward use of a single tool to analyze any or all of:

- Relative wealth (via scores)
- Absolute consumption expenditure (via poverty likelihoods and absolute poverty lines)
- Relative consumption expenditure (via poverty likelihoods and percentile-based poverty lines)

Unlike the scorecard, asset indexes only estimate relative wealth. Furthermore, the scorecard—unlike asset indexes—uses a straightforward, well-understood standard for socio-economic status whose definition is external to the tool itself (that is, consumption expenditure relative to a poverty line defined in monetary units).

In contrast, an asset index defines *poverty* in terms of its own questions and points, without calibration or reference to an external standard. This means that two asset indexes with different questions or different points—even if derived from the same data for a given country—imply two distinct definitions of *poverty*. In the same set-up, two scorecards would provide comparable estimates under a single definition of *poverty*.

---

[46] Percentiles are defined in terms of all people in Djibouti. For example, the all-Djibouti head-count poverty rate for the first-quintile (20th-percentile) poverty line is 20.0 percent (**Figure 10**).

[47] **Rutstein and Johnson**, 2004.

Annex 3 Scorecard construction

For Djibouti, about 70 candidate questions are prepared in these areas:

- Household composition (such as the number of household members)
- Education (such as whether the male head can read and write))
- Employment (such as the number of household members who work)
- Housing (such as the number of rooms or the main material of the floor)
- Ownership of consumer durables (such as TVs or cell phones)
- Agriculture (such as ownership of livestock)
- Location of residence (such as the region or arrondissement)

To facilitate the estimation of change over time, preference is given to questions that are more sensitive to changes in poverty. For example, the ownership of a cell phone is probably more responsive to changes in poverty than is the age of the head of the household).

The scorecard itself is built using 150% of the national poverty line and Logit regression on the construction sub-sample. Questions are selected based on both judgment and statistics.

The first step is to use Logit to build a draft scorecard for each candidate question. The power of each one-question draft scorecard to rank households by poverty status is assessed via the concentration index.[48]

[48] **Ravallion**, 2009.

One of the one-question draft scorecards is then selected based on:[49]

- Improvement in accuracy
- Acceptability to users in terms of:
  - Simplicity
  - Cost of collection
  - Concordance with:
    - Experience
    - Theory
    - Common sense
- Sensitivity to changes in consumption expenditure
- Variety among types of questions
- Applicability across regions and arrondissements
- Tendency to have a slow-changing relationship with poverty
- Relevance for distinguishing among people at the poorer end of the distribution of consumption expenditure
- Verifiability

A series of two-question draft scorecards are then built, each adding a second question to the one-question scorecard selected from the first stage. The best two-question draft scorecard is then selected, again using judgment to balance statistical accuracy with non-statistical criteria. These steps are repeated until the scorecard has 10 questions that work well together.

The last step is to transform the Logit coefficients into non-negative integers such that scores range from 0 to 100, with lower scores corresponding with greater poverty.

This algorithm is similar to common R^2-based stepwise least-squares regression. It differs from naïve stepwise in that the selection of questions considers both statistical[50] and non-statistical criteria. The use of non-statistical criteria can improve robustness through time and across non-nationally representative groups. It also helps to ensure that questions are straightforward, common-sense, inexpensive-to-collect, and acceptable to users.

---

[49] **Schreiner *et al.***, 2014.

[50] The statistical criterion is not the *p* values of coefficients but rather a question's contribution to the ranking of households by poverty status in the context of a scorecard with nine other questions.

The single scorecard here applies to all of Djibouti. Customizing poverty-assessment tools by urban/rural does not improve targeting accuracy much.[51] Segment-specific tools may improve the accuracy of estimates of poverty rates,[52] but:

- They run a greater risk of overfitting[53]
- Most of their benefit can be had in a single scorecard that includes a question that identifies the specific segment of interest (such as, in the case of Djibouti, the region or arrondissement of residence)[54]

[51] **Brown, Ravallion, and van de Walle**, 2018; **World Bank**, 2012; **Sharif**, 2009; **Schreiner**, 2006; **Schreiner**, 2005; **Narayan and Yoshida**, 2005; and **Grosh and Baker**, 1995.
[52] **Diamond *et al.***, 2016; **Tarozzi and Deaton**, 2009.
[53] **Haslett**, 2012.
[54] **Schreiner**, 2016b.

Annex 4 Estimates of poverty likelihoods

This annex tells how scores are converted into estimates of poverty likelihoods.

Scores are on an ordinal scale from 0 to 100. Higher scores signal less poverty, but not how much less. The ordered symbols used to represent scores are numbers, but those symbols do not stand for the normal cardinal numbers that you can do math on. For example, a score of 20 plus a score of 10 is not 30 of anything, just as the letter "A" plus the letter "B" is not the letter "C" (nor anything else).

To get cardinal units, a look-up table is used to convert scores to *poverty likelihoods*, that is, probabilities of having consumption expenditure below a poverty line. For the example of 100% of the national line, scores of 36−38 correspond with a poverty likelihood of 53.4 percent, and scores of 39−40 correspond with a poverty likelihood of 50.5 percent (**Figure 1**).

The poverty likelihood associated with a score varies by poverty line. For example, scores of 36−38 are associated with a likelihood of 53.4 percent for 100% of the national line but with a likelihood of 60.8 percent for the $3.20/day 2011 PPP line.

A4.1 Calibrating scores with poverty likelihoods

A given score is associated ("calibrated") with an estimated poverty likelihood that is defined as the share of people in the construction sub-sample who have the score and who live in households with per-capita or per-adult-equivalent consumption expenditure below a given poverty line.

For the example of 100% of the national line and a score of 36−38 (**Figure 22** below), there are 5,498 (normalized) households in the construction sample. Of these, 2,935 (normalized) have consumption expenditure below the poverty line. The estimated poverty likelihood associated with a score of 36−38 is then 53.4 percent, because 2,935 ÷ 5,498 ≈ 0.534 = 53.4 percent.

The same method is used to calibrate all scores with poverty likelihoods for all 16 poverty lines.[55]

[55] To ensure that likelihoods never increase as scores increase, likelihoods across adjacent scores may be non-parametrically smoothed before grouping scores into ranges. This preserves unbiasedness while preventing higher scores from being associated with higher likelihoods.

Figure 22: Estimation of poverty likelihoods (100% of national line)

Score	Households in range and < poverty line		All households in range		Poverty likelihood (%)
0–24	5,437	÷	5,953	=	91.3
25–28	3,718	÷	4,468	=	83.2
29–32	5,183	÷	6,389	=	81.1
33–35	3,209	÷	5,059	=	63.4
36–38	2,935	÷	5,498	=	53.4
39–40	2,133	÷	4,223	=	50.5
41–42	1,970	÷	3,992	=	49.4
43–44	1,468	÷	3,968	=	37.0
45–46	1,985	÷	5,498	=	36.1
47–48	1,276	÷	4,323	=	29.5
49–50	905	÷	4,019	=	22.5
51–52	594	÷	4,389	=	13.5
53–54	299	÷	4,859	=	6.2
55–56	181	÷	4,486	=	4.0
57–59	107	÷	6,174	=	1.7
60–62	89	÷	6,121	=	1.5
63–65	58	÷	4,415	=	1.3
66–69	26	÷	5,081	=	0.5
70–74	0	÷	4,700	=	0.0
75–100	0	÷	6,385	=	0.0

Number of all households normalized to sum to 100,000.

A4.2 Objectivity of estimates of poverty likelihoods

Even though scorecard questions are selected partly based on judgment related to non-statistical criteria, the calibration process produces estimates of poverty likelihoods that are objective, that is, derived from monetary poverty lines and from survey data on consumption expenditure.[56] The fact that some choices in scorecard construction are informed by judgment in no way impugns the objectivity of the estimated likelihoods; their objectivity depends on using data (and nothing else) in score calibration, not on using data (and nothing else) in scorecard construction.

A4.3 Why not use the Logit formula?

The scorecard is based on a Logit regression (**Annex 3**). This means that poverty likelihoods could be estimated not with a calibrated look-up table (**Figure 1**) but rather with the Logit formula of $2.718281828^{\beta X} \times (1 + 2.718281828^{\beta X})^{-1}$, where β is a vector of the Logit coefficients and X is a vector of a household's responses.

The scorecard uses the calibration approach is because the Logit formula looks scary. Program managers can understand poverty likelihoods defined as the share of people with a given score in the construction sample from Djibouti's 2017 EDAM who have consumption expenditure below a poverty line. A calibrated look-up table also allows analysts to convert scores to likelihoods without any math at all. This calibration approach can also improve accuracy, especially with large samples.

[56] The calibrated likelihoods would be objective even if scorecard construction did not use any data at all. In fact, objective scorecards of proven accuracy are often constructed using only expert judgment (**Caire**, 2004; **Schreiner *et al.***, 2014).

Annex 5 Error and margins of error

This annex discusses the scorecard's estimation error for head-count poverty rates in a single time period, as well as margins of error for all estimates.

A5.1 Estimation errors

A5.1.1 What is estimation error?

For example, the estimation error of Djibouti's scorecard for estimates of head-count poverty rates in a single time period by 100% of the national poverty line is −1.5 percentage points (**Figure 2**).

An unadjusted estimate can usually be improved—that is, moved closer to the true value in the population—by subtracting off the known estimation error. For example, if the unadjusted estimate is 87.4 percent and the estimation error is −1.5 percentage points, then an improved estimate is 87.4 − (−1.5) = 88.9 percent.

A5.1.2 What estimation errors are reported for the Djibouti scorecard?

Estimation errors are reported for estimates of head-count poverty rates in a single time period for all 16 poverty lines supported by the scorecard for Djibouti.

The estimation errors are derived *out-of-sample*. This means that the scorecard (made from the construction sample from the 2017 EDAM, **Annex 1**) is tested with repeated sub-samples of households from the validation sample that were not used to construct the scorecard. The estimation error is the average of the differences between scorecard estimates and observed poverty rates across these repeated sub-samples.

There is no data now on expenditure-based poverty in the future, so it is impossible to report estimation errors for annual net changes in head-count poverty rates across two time periods. The scorecard cannot be tested *out-of-time* because it is both constructed and validated with data from a single time period (2017).

In practice, the scorecard—like all poverty-assessment tools—is always applied both out-of-sample and out-of-time. Being out-of-sample violates the assumption that the scorecard is applied to a sample from the same population whose data was used to construct the scorecard. Being out-of-time violates the assumption that the relationships between poverty and scorecard questions are the same as in the population whose data was used to construct the scorecard.

A5.1.3 How to estimate estimation errors

Given the scorecard's standard assumptions, an unbiased estimator of *estimation error* is the average of differences between scorecard estimates and observed values in repeated sub-samples from the validation sample.[57]

It is possible to compare estimated and observed poverty rates because the 2017 EDAM records actual (not estimated) expenditure-based poverty status for households in the validation sample. The observed poverty likelihood in the 2017 EDAM is 100 percent for poor households and 0 percent for non-poor households. For a given poverty line, the observed (not estimated) head-count poverty rate is the household-size-weighted average of observed poverty likelihoods.

The scorecard can also be applied to the same validation sub-sample (ignoring that actual poverty status is observed) to estimate the poverty rate as the household-size-weighted average of estimated poverty likelihoods (**Section 3**).

The scorecard's error in a given validation sub-sample is then the difference between the scorecard's estimate versus the observed value.

Different sub-samples from the validation sample result in different errors. The estimate of the scorecard's general *estimation error* is the average of these errors across many sub-samples.[58] In turn, the scorecard estimate's margin of error reflects the extent of the spread of the distribution of all the sub-samples' errors around their average.[59]

A5.1.4 Estimation errors for estimates of poverty rates in one time period

The first line in **Figure 2** ("Estimation error") presents estimation errors for estimates of poverty rates in one time period for Djibouti's 16 poverty lines.

---

[57] This is the *bootstrap approach*. The average of estimates from repeated samples from the validation sample is an unbiased estimator of the true value in the population of Djibouti overall. The population's true value is taken as the value in the 2017 EDAM (even though the EDAM is itself only a sample).

[58] Households in a sub-sample are drawn *with replacement*; each draw is from the full pool, including households that have already been drawn. Thus, a given household may appear in a given sub-sample once, more than once, or not at all.

[59] **Schreiner** (2021) discusses the derivation of errors and estimation error.

A5.2 Margins of error

A5.2.1 What are margins of error?

Like any statistic, a scorecard estimate depends on a particular sample from a population. Because samples are drawn at random, each sample is different, and different samples give different scorecard estimates. Scorecard estimates are *unbiased*—under the standard assumptions—because the average of scorecard estimates across many repeated samples is the same as the single true value in the population.

In any single sample, however, unusual luck may push an estimate for that sample far from the true value in the population. Larger samples provide more chances for luck to even out, so large errors are less likely in larger samples.[60]

For a given estimate, sample size, and confidence level, the *margin of error* is the range of true population values that is (in some degree) consistent with the estimate.

A margin of error has two parts:

- The margin of error itself (such as ±2.0 percentage points). This range is centered on the estimate
- A confidence level (such as 90 percent) that the true value falls within the margin of error

All else constant, narrower margins of error or higher confidence levels mean that it is more likely that the sample-based estimate is closer to the true population value.

[60] When flipping a fair (unbiased) coin, the true probability of "heads" is 50 percent. *Unbiasedness* means that the average of the share of "heads" across many samples will be close to 50 percent. In a single sample of 10 tosses, however, the chances of getting at least six "heads" (60 percent of the 10 tosses, with an error of at least 10 percentage points) is about 37 percent. In a single sample of 100 tosses, the chances of such a large error is smaller (about 3 percent). Larger samples reduce the risk that estimates will be far from true values.

To illustrate, suppose that the adjusted estimate of the head-count poverty rate for 100% of the national line is 88.9 percent and that the sample size is n = 1,024. Given 90-percent confidence,[61] the margin of error is ±2.4 percentage points (**Figure 2**). Absent other sources of error and given the scorecard's standard assumptions, this means that there is a 90-percent chance that the true population value is in the range from 88.9 − 2.4 = 86.5 percent to 88.9 + 2.4 = 91.3 percent, with the most-likely true value being the center of the range (the 88.9-percent estimate).

Said another way, "With 90-percent confidence, the estimate has a margin of error from 86.5 to 91.3 percent." This means that the true population value has a:

- 5-percent chance of being less than 86.5 percent
- 90-percent chance of being between 86.5 and 91.3 percent
- 5-percent chance of being greater than 91.3 percent

A5.2.2 Why do margins of error matter?

Managers should put more weight on estimates with narrower margins of error.

As a hypothetical example, a pro-poor program in Djibouti probably is indeed pro-poor if the scorecard estimate of the poverty rate for in-coming participants by 100% of the national poverty line with 80-percent confidence is 50.0 percent with a margin of error of ±5.0 percentage points, that is, from 45.0 to 55.0 percent. The estimate and its margin of error suggest that the true poverty rate of in-coming participants is unlikely to be less than or about the same as the all-Djibouti rate for this line of 35.8 percent (**Figure 10**).

If, however, the margin of error were ±15.0 percentage points (that is, from 35.0 to 65.0 percent), then there is a non-negligible chance that the poverty rate of in-coming participants is less than or about the same as that for Djibouti overall (35.8 percent) and thus that the program may not actually be pro-poor.

So far, almost all analyses of scorecard estimates have ignored margins of error. This deficient practice increases the risk of bad decisions. Do not make this mistake.

[61] Most real-world decisions are made with much less than 90-percent confidence.

A5.2.3 Margins of error for estimates of poverty rates in one time period for the Djibouti scorecard

For sample sizes of n = 1,024 and 90-percent confidence and across all supported poverty lines, the margins of error for estimates of head-count poverty rates in a single time period for the Djibouti scorecard are ±2.5 percentage points or smaller (**Figure 2**). Given the scorecard's standard assumptions, this means that in 90 of 100 samples of this size, the true population value is within ±2.5 percentage points or less of the error-adjusted estimate.

A5.2.4 How to calculate margins of error

The spreadsheet-based **ProveIt!™ reporting app** calculates margins of error for all scorecard estimates discussed here. Analysts may also use the formulas below.[62]

A5.2.5 Formula for margins of error for estimates of head-count poverty rates in a single time period

All formulas for margins of error involve the following elements:

±c is the margin of error as a proportion (*e.g.*, ±0.020 for ±2.0 percentage points),

z is from the Normal distribution and is $\begin{cases} 1.04 \text{ for confidence levels of 70 percent} \\ 1.28 \text{ for confidence levels of 80 percent} \\ 1.64 \text{ for confidence levels of 90 percent} \end{cases}$,

σ is the standard error of the estimated poverty rate, that is, $\sqrt{\frac{\hat{p} \cdot (1-\hat{p})}{n}} \cdot \varphi$,

$\hat{p}$ is the estimated poverty rate as a proportion,

φ is the finite population correction factor $\sqrt{\frac{N-n}{N-1}}$,

N is the population size in terms of households (not members of households),

n is the sample size (in terms of interviewed households,
not members of interviewed households), and

α is an adjustment factor specific to the scorecard, estimator, and poverty line.

[62] **Schreiner** (2021) discusses the derivation of the formulas.

Given a confidence level that corresponds with z, a sample-based estimate $\hat{p}$, a population N, a sample n, and an adjustment factor α for a specific poverty line from **Figure 2**, the formula[63] for the margin of error $\pm c$ is $\pm z \cdot \alpha \cdot \sqrt{\frac{\hat{p} \cdot (1-\hat{p})}{n}} \cdot \sqrt{\frac{N-n}{N-1}}$.

To illustrate, Djibouti's 2017 EDAM gives a direct-measure head-count poverty rate for 100% of the national line of $\hat{p}$ = 35.8 percent (**Figure 10**). The adjustment factor α is 1.00 by definition because $\hat{p}$ is a direct-measure estimate, not an indirect-scorecard estimate.[64] Djibouti in 2017 had a population of households (not people) of N = 104,162, and the EDAM sample size was n = 4,474. Given a desired confidence level of 90 percent, z is 1.64. The margin of error $\pm c$ is then about ±1.1 percentage points:

$$\pm z \cdot \alpha \cdot \sqrt{\frac{\hat{p} \cdot (1-\hat{p})}{n}} \cdot \sqrt{\frac{N-n}{N-1}} = \pm 1.64 \cdot 1.00 \cdot \sqrt{\frac{0.358 \cdot (1-0.358)}{4{,}474}} \cdot \sqrt{\frac{104{,}162-4{,}474}{104{,}162-1}}.$$

This implies a 90-percent chance that Djibouti's true head-count poverty rate for 100% of the national line in 2017 is in the range from 35.8 − 1.1 = 34.7 percent to 35.8 + 1.1 = 36.9 percent.

A5.2.6 Margins of error for estimates of numbers of poor people in a single time period

The lower (upper) limit of the margin of error for an estimate of numbers of poor people is the number of people in participating households, multiplied by the lower (upper) limit of the margin of error of the poverty-rate estimate.

[63] This formula ignores how sampling variability affects the derivation of the scorecard. It also ignores that household size varies and that larger households are more likely to have higher poverty likelihoods. This understates the margin of error.
[64] For scorecard estimates, α for a given poverty line is found in **Figure 2**.

To illustrate, the baseline example in **Section 33.1** has an estimated poverty rate of 88.9 percent. With 70-percent confidence, the margin of error is about ±26.1 percentage points,[65] or from 88.9 − 26.1 = 62.8 percent to 88.9 + 26.1 = 115.0 percent ≈ 100.0 percent (because poverty rates cannot exceed 100 percent). The margin of error is huge because the sample size of n = 2 interviewed households is very small.[66]

The estimated number of people in participating households in the example in **Section 33.2** is 6,500,[67] so the lower limit of the 70-percent margin of error for the estimated number of poor people is 6,500 · 0.628 = 4,082. The upper limit is 6,500 · 1.000 = 6,500. This example estimate—based as it is on a sample of two households—is better understood not as "about 90 percent of people in participating households are poor" but rather as "at least 60 percent are poor".

A5.2.7 Margins of error for estimates of the annual net change in head-count poverty rates across two periods for one sample, scored twice

In this case, the formula for the margin of error ±c is:

$$\pm\frac{z\cdot\alpha}{y}\cdot\sqrt{\frac{\hat{p}_{up}\cdot(1-\hat{p}_{up})+\hat{p}_{down}\cdot(1-\hat{p}_{down})+2\cdot\hat{p}_{up}\cdot\hat{p}_{down}}{n}}\cdot\sqrt{\frac{N-n}{N-1}}\text{, where}$$

- z, α, N, and n are defined as above
- $\hat{p}_{up}$ is the share of members of sampled households that rise above the poverty line from below
- $\hat{p}_{down}$ is the share of members of sampled households that fall below the poverty line from above
- y is the household-size-weighted average of years between interviews

---

[65] The example in **Section 33.2** has N = 1,000, n = 2, and α = 1.13 (**Figure 2**). For 70-percent confidence, z = 1.04. The margin of error ±c for the head-count poverty rate estimate is then $\pm 1.04\cdot 1.13\cdot\sqrt{\frac{0.889\cdot(1-0.889)}{2}}\cdot\sqrt{\frac{1{,}000-2}{1{,}000-1}}$ ≈ ±0.261.

[66] Yet the formulas for margin of error still apply, and the estimator is still unbiased.

[67] The formula for margin of error for the estimated number of poor people ignores that the estimated number of people in participating households has its own margin of error. This understates the margin of error.

Illustrating with the earlier example of one sample scored twice (**Section 3.3.1**), $\hat{p}_{up}$ is the share of household members estimated to rise above a poverty line from below. This is the absolute value of the sum of the estimated *negative* changes in the number of members in poor households (from rows 3 and 4 of column M in **Figure 11**, here | −0.608 + −0.974 | = 1.582, divided by the sum across all sampled households of each household's average household size across baseline and follow-up of 6.5 + 6.5 = 13.0 (from row 5, columns E and F). Thus, $\hat{p}_{up}$ = 1.582 ÷ 13.0 = 0.122.

In turn, $\hat{p}_{down}$ is the share of household members estimated to fall below a poverty line from above. This is the sum of the estimated *positive* net changes in the number of members in poor households (from rows 3 and 4 of column M in **Figure 11)**, which is (+0.00) + (+0.00) = +0.000 (because the estimated poverty likelihood did not increase for any households). Dividing this by the sum across all sampled households of each household's average household size across baseline and follow-up (6.5 + 6.5 = 13.0) gives $\hat{p}_{down}$ = 0.000 ÷ 13.0 ≈ 0.000.[68]

The household-size-weighted average of the number of years between interviews *y* is 3.05 (from row 9, column M in **Figure 11)**.

With sample size *n* = 2 interviewed households, population *N* of 1,000 households, confidence level of 70 percent (*z* = 1.04), and the α adjustment factor for this estimator (regardless of poverty line) of 1.14,[69] the margin of error ±*c* is about ±0.090 ≈

$$\pm\frac{1.04\cdot 1.14}{3.05}\cdot\sqrt{\frac{0.122\cdot(1-0.122)+0.000\cdot(1-0.000)+2\cdot 0.122\cdot 0.000}{2}}\cdot\sqrt{\frac{1{,}000-2}{1{,}000-1}}.$$

The example's estimated net annual poverty-rate change is about −4.0 percentage points (**Figure 11**), so the 70-percent margin of error is from −4.0 − 9.0 = −13.0 to −4.0 + 9.0 = +5.0 percentage points. The margin of error shows that—due to the tiny sample of *n* = 2—this estimate is uninformative; the true net change in the population could very well be strongly negative, close to zero, or strongly positive.

---

[68] $\hat{p}_{down} - \hat{p}_{up}$ is the estimated net poverty-rate change. In this example, $\hat{p}_{down}$ ≈ 0.000 and $\hat{p}_{up}$ = 0.122, so 0.000 − 0.122 ≈ −0.122, which is the estimated 12.2 percentage-point decrease (improvement) in the poverty rate in **Figure 11**.

[69] **Schreiner**, 2021.

This example shows why margins of error are useful. Without them, program managers might believe that there was evidence that poverty rates decreased by about 4.0 percentage points per year even though the data in this sample is also consistent with widely different rates and directions of change.

A5.2.8 Margins of error for estimates of the annual net change in the number of poor people across two periods for one sample, scored twice

The lower (upper) limit of the margin of error for an estimate of annual net change in the number of poor people for one sample, scored twice is the average number of people in participating households from baseline to follow-up, multiplied by the lower (upper) limit of the margin of error of the estimated annual net change in the poverty rate.

To illustrate with the example in **Section** **3.3.13.3.2** for one sample scored twice, the estimated annual net change in the poverty rate is about −4.0 percentage points. As just shown, the tiny sample size of $n = 2$ means that the 70-percent margin of error runs from −13.0 to +5.0 percentage points.

The estimated average number of on-going participating people per year is 5,525.[70] Thus, the lower limit of the 70-percent margin of error for the estimated annual net change in the number of poor people is $5{,}525 \cdot (-0.130) \approx -718$ (a net decrease in poor people), and the upper limit is $5{,}525 \cdot (+0.050) \approx +276$ (a net increase in poor people). The small sample leads to a large margin of error, so the estimate is not likely to be useful because it is consistent with a true reduction, a true increase, or no true change at all.

[70] The formula for margin of error for the estimated number of poor people ignores that the estimated number of people in participating households has its own margin of error. This understates the margin of error.

A5.2.9 Margins of error for estimates of the annual net change in head-count poverty rates across two periods for two independent samples

The formula for the margin of error ±*c* is $\pm\frac{z\cdot\alpha}{y}\cdot\sqrt{\frac{2\cdot\hat{p}\cdot(1-\hat{p})}{n}}\cdot\sqrt{\frac{N-n}{N-1}}$.

with *z*, α, *y*, $\hat{p}$ and *N* defined as above. Then *n* households are sampled and interviewed at baseline, and another *n* households are sampled and interviewed at follow-up.

Illustrating with the example for two independent samples in **Section 3.3.4**:

- $z = 1.04$, assuming a desired confidence level of 70 percent
- $\alpha = 1.10$, the adjustment factor (regardless of poverty line) for this estimator[71]
- $y = 2.76$, the years between the average interview at baseline and follow-up
- $\hat{p} = 0.874$, the unadjusted) estimate of the poverty rate at baseline
- $N = 850$, the average number of households across baseline (1,000) and follow-up (700)
- $n = 2$, the sample size in both baseline and follow-up

The margin of error ±*c* is $\pm 0.137 \approx \pm\frac{1.04\cdot1.10}{2.76}\cdot\sqrt{\frac{2\cdot0.874\cdot(1-0.874)}{2}}\cdot\sqrt{\frac{850-2}{850-1}}$.

The example's estimated net annual poverty-rate change is −8.7 percentage points (**Figure 12**). Thus, the 70-percent margin of error is from −8.7 − 13.7 = −22.4 percentage points to −8.7 + 13.7 = +5.0 percentage points. The tiny sample is consistent with a true value in the population that is strongly negative, close to zero, or strongly positive. This again shows why margins of error matter.

[71] **Schreiner**, 2021.

A5.2.10 Margins of error for estimates of the annual net change in the number of poor people across two periods for two independent samples

The lower (upper) limit of the margin of error for an estimate of annual net change in the number of poor people for two independent samples is the average number of people in participating households from baseline to follow-up, multiplied by the lower (upper) limit of the margin of error of the estimated annual net change in the poverty rate.

To illustrate, the example in **Section 3.3.4** for two independent samples estimates the annual net change in the poverty rate as −8.7 percentage points. As just shown, the 70-percent margin of error runs from −22.4 to +5.0 percentage points.

The estimated average number of on-going participating people is 5,175.[72] Thus, the lower limit of the 70-percent margin of error for the estimated annual net change in the number of poor people per year is 5,175 · (−0.224) ≈ −1,159 (a net decrease in poor people), and the upper limit is 5,175 · (+0.050) ≈ +259 (a net increase in poor people). The margin of error again shows that the estimate is not very informative of the true value in the population.

[72] The formula for margin of error for the estimated number of poor people ignores that the estimated number of people in participating households has its own margin of error. This understates the margin of error.

Annex 6 Formulas for sample size

Before drawing a sample of households to interview, the formulas here can be used to calculate the sample size that corresponds to a program's:

- Desired margin of error for the eventual scorecard estimate, and
- Desired confidence level for the margin of error, and
- Pre-estimation guess of the true population value to be estimated

These formulas may or may not be useful, for several reasons.

First, programs sometimes collect scorecard data but then fail to report and analyze it. In such cases, the entire project is a waste, so there is no point in worrying about sample size. This is why programs must plan and budget for reporting and analysis. If the remaining budget (after planning for reporting and analysis) will not cover at least 1,000 interviews, then ignore the formulas below and do as many interviews as the budget allows.

Second, both statistical sample size and psychological sample size matter. On the one hand, samples smaller than n = 300 often seem too small. On the other hand, samples of at least n = 1,000 usually seem large enough.

Third, calculating an optimal sample size makes sense only if a program:

- Has reason to desire a particular margin of error or level of confidence[73]
- Plans to report and analyze margins of error (as already mentioned)

If margins of error are not understood or will not be reported and analyzed, then just interview as many participating households as the budget allows.

Fourth, sample-size calculations are sometimes unneeded. For example, using the scorecard for segmenting requires interviewing all relevant participants. Likewise, doing a basic check on the fulfillment of a pro-poor mission may be less costly if all in-coming participants are scored as a routine step of the in-take process rather than repeatedly deciding at the moment whether to score a given enrollee.

[73] Academic conventions for levels of confidence, applied to business, often imply unnecessarily large samples.

In sum, go ahead with the formulas below if you:

- Reserve resources for reporting and analysis, and
- Understand margins of error and will report and analyze them, and
- Plan to estimate net changes in poverty over time, and
- Can afford at least 1,000 interviews at both baseline and follow-up

Otherwise:

- If checking a pro-poor mission, then score all in-coming participants at in-take
- If segmenting by poverty, then score all relevant participants
- If estimating changes in poverty, then score as many participants as the budget allows

A6.1 Sample-size formula for estimates of head-count poverty rates in a single time period

In this case, the formula for the sample size n (the number of participating households to be interviewed) is $n = N \cdot \left(\frac{z^2 \cdot \alpha^2 \cdot \tilde{p} \cdot (1-\tilde{p})}{z^2 \cdot \alpha^2 \cdot \tilde{p} \cdot (1-\tilde{p}) + c^2 \cdot (N-1)} \right)$,

where n, c, z, α, and N are defined as in **Annex 5**, and $\tilde{p}$ is a before-estimation guess for the poverty rate to be estimated.[74]

The illustration below of the calculation of the sample size n uses these values:

- The population of participating households is N = 10,000
- The desired confidence level for the margin of error is 80 percent, so z = 1.28
- The poverty line is 100% of the national line, so α = 1.13 (**Figure 2**)
- The pre-estimation expected poverty rate is the all-Djibouti rate for 100% of the national line in 2017, so $\tilde{p}$ = 35.8 percent = 0.385 (**Figure 10**)
- The desired margin of error $\pm c$ = ±3.0 percentage points = ±0.030

Given these hypothetical values,

$$n = 10{,}000 \cdot \left(\frac{1.28^2 \cdot 1.13^2 \cdot 0.358 \cdot (1-0.358)}{1.28^2 \cdot 1.13^2 \cdot 0.358 \cdot (1-0.358) + 0.03^2 \cdot (10{,}000-1)} \right) \approx 508.$$

[74] If the population N is "large" relative to the expected sample size n, then the formula can be taken as $n = \left(\frac{\alpha \cdot z}{c} \right)^2 \cdot \tilde{p} \cdot (1-\tilde{p})$.

A6.2 Sample-size formula for estimates of annual net changes in head-count poverty rates across two time periods with one sample scored twice

In this case, *n* households are interviewed at baseline, and those same *n* households are interviewed again at follow-up. The formula for *n* is:[75]

$$2 \cdot \left(\frac{z \cdot \alpha}{c} \right)^2 \cdot [-0.01 + 0.016 \cdot y + 0.56 \cdot p_{\text{pre-baseline}} \cdot (1 - p_{\text{pre-baseline}})] \cdot \sqrt{\frac{N - n}{N - 1}},$$

where *n*, α, *z*, *c*, and *N* are defined as above, *y* is the number of years between baseline and follow-up, and $p_{\text{pre-baseline}}$ is the population's expected head-count poverty rate prior to the baseline interviews.

The illustration below for this formula uses the following values:

- The poverty line is 100% of the national line
- The desired confidence level for the margin of error is 80 percent, so $z = 1.28$
- $\alpha = 1.14$ (regardless of the scorecard or poverty line[76]
- The desired margin of error $\pm c = \pm 3.0$ percentage points = ±0.030
- The number of years between baseline and follow-up is $y = 3$
- The pre-estimation expected pre-baseline poverty rate is the all-Djibouti rate for 100% of the national line: $p_{\text{pre-baseline}} = 35.8$ percent = 0.358 (**Figure 10**)
- The population of participating households is $N = 10{,}000$

Assuming *N* is large relative to *n* so that $\sqrt{\frac{N - n}{N - 1}} \approx 1$, then the baseline sample size *n* is $2 \cdot \left(\frac{1.28 \cdot 1.14}{0.03} \right)^2 \cdot [-0.01 + 0.016 \cdot 3 + 0.56 \cdot 0.358 \cdot (1 - 0.358)] \cdot 1 \approx 789.$

The follow-up sample size is also 789.

[75] **Schreiner**, 2021.
[76] **Schreiner**, 2021.

A6.3 Sample-size formula for estimates of annual net changes in head-count poverty rates across two time periods with two independent samples

This formula is two (2), multiplied by the formula for sample size for an estimate at a point in time. If n and $\tilde{p}$ are the same at both baseline and follow-up, then $n = 2 \cdot N \cdot \left(\frac{z^2 \cdot \alpha^2 \cdot \tilde{p} \cdot (1-\tilde{p})}{z^2 \cdot \alpha^2 \cdot \tilde{p} \cdot (1-\tilde{p}) + c^2 \cdot (N-1)} \right)$.[77]

There are n interviews at baseline, and another n interviews at follow-up. For this estimator and regardless of the scorecard or poverty line, $\alpha = 1.10$.[78]

To illustrate with the same hypothetical values as in the example just above (except that $\alpha = 1.10$), the sample size at baseline n is:

$$2 \cdot 10{,}000 \cdot \left(\frac{1.28^2 \cdot 1.10^2 \cdot 0.358 \cdot (1-0.358)}{1.28^2 \cdot 1.10^2 \cdot 0.358 \cdot (1-0.358) + 0.03^2 \cdot (10{,}000-1)} \right) \approx 964.$$

The sample size at follow-up is also $n = 964$.

[77] If the N is large relative to n, then the formula is about $n = 2 \cdot \left(\frac{\alpha \cdot z}{c} \right)^2 \cdot \tilde{p} \cdot (1-\tilde{p})$.

[78] **Schreiner**, 2021.

References

Adams, Niall M.; and David J. Hand. (2000) "Improving the Practice of Classifier Performance Assessment", *Neural Computation*, Vol. 12, pp. 305−311, **request**.

Brown, Caitlin; Ravallion, Martin; and Dominique van de Walle. (2018) "A Poor Means Test? Econometric Targeting in Africa", *Journal of Development Economics*, Vol. 134, pp. 109−124, **link**.

Caire, Dean. (2004) "Building Credit Scorecards for Small-Business Lending in Developing Markets", **link**.

_____; and Mark Schreiner. (2012) "Cross-Tab Weighting for Credit Scorecards in Developing Markets", **link**.

Coady, David; Grosh, Margaret; and John Hoddinott. (2004) *Targeting of Transfers in Developing Countries*, **link**.

Dawes, Robyn M. (1979) "The Robust Beauty of Improper Linear Models in Decision Making", *American Psychologist*, Vol. 34, No. 7, pp. 571−582, **link**.

Deaton, Angus; and Salman Zaidi. (2002) "Guidelines for Constructing Consumption Aggregates for Welfare Analysis", World Bank LSMS Working Paper No. 135, **link**.

Diamond, Alexis; Gill, Michael; Rebolledo Dellepiane, Miguel Angel; Skoufias, Emmanuel; Vinha, Katja; and Yiqing Xu. (2016) "Estimating Poverty Rates in Target Populations: An Assessment of the Simple Poverty Scorecard and Alternative Approaches", World Bank Policy Research Working Paper No. 7793, **link**.

Direction de la Statistique et des Études Démographiques. (2018). « Résultats de la Quatrième Enquête Djiboutienne Auprès des Ménages pour les Indicateurs Sociaux (EDAM-IS) », **link**.

Dupriez, Olivier. (2018) "An Empirical Comparison of Machine-Learning Classification Algorithms", **link**.

Ferreira, Francisco H.G.; Chen, Shaohua; Dabalen, Andrew L.; Dikhanov, Yuri; Hamadeh, Nada; Jolliffe, Dean; Narayan, Ambar; Prydz, Espen Beer; Revenga, Ana; Sangraula, Prem; Serajuddin, Umar; and Nobuo Yoshida. (2016) "A Global Count of the Extreme Poor in 2012: Data Issues, Methodology, and Initial Results", *Journal of Economic Inequality*, Vol. 14, pp. 141−172, **link**.

Grosh, Margaret; and Judy L. Baker. (1995) “Proxy-Means Tests for Targeting Social Programs: Simulations and Speculation”, World Bank LSMS Working Paper No. 118, **link**.

Hand, David J. (2006) “Classifier Technology and the Illusion of Progress”, *Statistical Science*, Vol. 22, No. 1, pp. 1−15, **link**.

Haslett, Stephen. (2012) “Practical Guidelines for the Design and Analysis of Sample Surveys for Small-Area Estimation”, *Journal of the Indian Society of Agricultural Statistics*, Vol. 66, No. 1, pp. 203−212, **request**.

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, **request**.

IRIS Center. (2007) “Manual for the Implementation of USAID Poverty-Assessment Tools”, **link**.

Jolliffe, Dean; and Espen Beer Prydz. (2016) “Estimating International Poverty Lines from Comparable National Thresholds”, *Journal of Economic Inequality*, Vol. 14, pp. 185−198, **request**.

Kolesar, Peter; and Janet L. Showers. (1985) “A Robust Credit-Screening Model Using Categorical Data”, *Management Science*, Vol. 31, No. 2, pp. 124−133, **link**.

Lovie, Alexander D.; and Patricia Lovie. (1986) “The Flat-Maximum Effect and Linear Scoring Models for Prediction”, *Journal of Forecasting*, Vol. 5, pp. 159−168, **link**.

Matul, Michal; and Sean Kline. (2003) “Scoring Change: Prizma’s Approach to Assessing Poverty”, MFC Spotlight Note No. 4, **link**.

Myers, James H.; and Edward W. Forgy. (1963) “The Development of Numerical Credit-Evaluation Systems”, *Journal of the American Statistical Association*, Vol. 58, No. 303, pp. 779−806, **link**.

Narayan, Ambar; and Nobuo Yoshida. (2005) “Proxy-Means Tests for Targeting Welfare Benefits in Sri Lanka”, World Bank Report No. SASPR−7, **link**.

Navajas, Sergio; Schreiner, Mark; Meyer, Richard L.; Gonzalez-Vega, Claudio; and Jorge Rodriguez-Meza. (2000) “Microfinance and the Poorest of the Poor: Theory and Evidence from Bolivia”, *World Development*, Vol. 28, No. 2, pp. 333−346, **link**.

Ravallion, Martin. (2009) “How Relevant Is Targeting to the Success of an Anti-Poverty Program?”, *World Bank Research Observer*, Vol. 24, No. 2, pp. 205−231, **request**.

_____. (1998) “Poverty Lines in Theory and Practice”, World Bank LSMS Working Paper No. 133, **link**.

Rutstein, Shea Oscar; and Kiersten Johnson. (2004) “The DHS Wealth Index”, DHS Comparative Reports No. 6, **link**.

Schreiner, Mark. (2021) “Scorocs Simple Poverty Scorecard Tool: Uganda”, **link**.

_____. (2016a) “Popular Science”, **link**.

_____. (2016b) “Simple Poverty Scorecard Tool: Indonesia (Jawa Timur and Nusa Tengarra Timur)”, **request**.

_____. (2015) “There’s No Place Like Home? How the Interview Method Affects Results with the Progress out of Poverty Index®”, **link**.

_____. (2014) “The Process of Poverty-Scoring Analysis”, **link**.

_____. (2013) “Scorocs Simple Poverty Scorecard Tool: Bangladesh”, **link**.

_____. (2012) “An Expert-Based Poverty Scorecard for Rural China”, **link**.

_____. (2006) “Is One Simple Poverty Scorecard Enough for India?”, **link**.

_____. (2005) “La Herramienta del Índice de Calificación de la Pobreza™ de Scorocs: México”, **link**.

_____. (2002) *Scoring: The Next Breakthrough in Microfinance?* CGAP Occasional Paper No. 7, **link**.

_____. (1997) *A Framework for the Analysis of the Performance and Sustainability of Subsidized Microfinance Organizations with Application to BancoSol of Bolivia and the Grameen Bank of Bangladesh*, **link**.

_____; Matul, Michal; Pawlak, Ewa; and Sean Kline. (2014) “Poverty Scoring: Lessons from a Microlender in Bosnia-Herzegovina”, *Poverty and Public Policy*, Vol. 6, No. 4, pp. 407−428, **link**.

Sharif, Iffath Anwar. (2009) “Building a Targeting System for Bangladesh Based on Proxy-Means Testing”, World Bank Social Protection Discussion Paper No. 0914, **link**.

Stillwell, William G.; Barron, F. Hutton; and Ward Edwards. (1983) "Evaluating Credit Applications: A Validation of Multi-Attribute Utility-Weight Elicitation Techniques", *Organizational Behavior and Human Performance*, Vol. 32, pp. 87−108, **request**.

Tarozzi, Alessandro; and Angus Deaton. (2009) "Using Census and Survey Data to Estimate Poverty and Inequality for Small Areas", *Review of Economics and Statistics*, Vol. 91, No. 4, pp. 773−792, **link**.

Toohig, Jeff. (2008) "PPI® Pilot Training Guide", **link**.

Wainer, Howard. (1976) "Estimating Coefficients in Linear Models: It Don't Make No Nevermind", *Psychological Bulletin*, Vol. 83, pp. 223−227, **link**.

World Bank. (2020) *Purchasing Power Parities and the Size of World Economies: Results from the 2017 International Comparison Program*, **link**.

_____. (2013) "Shared Prosperity: A New Goal for a Changing World", **link**.

_____. (2012) *Targeting Poor and Vulnerable Households in Indonesia*, **link**.