Simple Poverty Scorecard® Tool
Bolivia

Mark Schreiner

20 February 2023

Consultar este documento y una **aplicación de recopilación de datos** en Castellano en **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 Bolivia 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

Version note

The new scorecard for Bolivia is based on data from **2019**. It replaces old scorecards for Bolivia based on data from **2013**, **2007**, and **2002**. Because Bolivia **changed its definition of *poverty*** in 2019, it is not possible to estimate changes with a baseline from an old scorecard and a follow-up from the new scorecard.

Acknowledgements

This work was funded by Chalice. Data are from Bolivia's *Instituto Nacional de la Estadística*. Thanks go to MaryAnne McKinnon.

Scorocs® Simple Poverty Scorecard® Poverty Assessment Tool

			Full name	Identifier
Interview ID:				
Interview date:		Participant of record:		
Country:	BOL	Service agent:		
Scorecard:	004	Service point:		
Sampling weight:		Number of household members:		

Question	Response	Points
1. In which department does the household live? (*record without asking*)	A. Chuquisaca	0
	B. La Paz	4
	C. Tarija, or Oruro	6
	D. Cochabamba, or Potosí	7
	E. Pando, or Beni	11
	F. Santa Cruz	13
2. How many members does the household have? (*from **Back-page Worksheet***)	A. Six or more	0
	B. Five	4
	C. Four	10
	D. Three	16
	E. Two	23
	F. One	28
3. How many household members work in their main occupation as wage/salary employees, domestic servants, or business owners who draw a salary? (*from **Back-page Worksheet***)	A. None	0
	B. One	9
	C. Two or more	17
4. How many household members in their main occupation are self-employed, business owners who do not draw a salary, or members of a producer's cooperative? (*from **Back-page Worksheet***)	A. None	0
	B. One	4
	C. Two or more	12
5. Is the female head (or spouse of the male head) covered by medical insurance?	A. No	0
	B. Yes	2
	C. No female head (nor spouse of the male head)	4
6. How many rooms does the household occupy, not counting bathrooms, kitchen, laundry rooms, garages, or rooms used for storage or business?	A. One	0
	B. Two or three	3
	C. Four or more	5
7. Is firewood, dung, manure, or llama pellets the main fuel or source of energy used for cooking?	A. Yes	0
	B. No	4
8. Does the household possess and use a frigerator or freezer?	A. No	0
	B. Yes	2
9. Does the household possess and use an old-style tube TV or one or more flat-screen TVs (plasma, LCD, LED)?	A. No	0
	B. Only tube	2
	C. One flat screen (regardless of tube)	6
	D. Two or more flat screens (regardless of tube)	9
10. How many household members have a cell phone available for their own personal (non-business) use?	A. None	0
	B. One	2
	C. Two	3
	D. Three or more	6

Score:

Back-page Worksheet

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 for the participant of record (who may differ from the respondent), for the service agent of the participant of record (who may differ from you the enumerator), and for the service point that the participant of record uses (if any and if known). Without asking the respondent, circle the response to the first scorecard question based on the department where the household lives.

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 spouse/conjugal partner (if there is one). A household is a unit of people (regardless of blood or marital relationship) who usually live in the same residence and who together meet their basic needs from a shared budget. That is, household members share expenses, regardless of who provides the resources to meet those expenses. Household members include those who currently live with the household and whose actual or expected total stay is at least 3 months, as well as those who are temporarily absent (if their total expected absence is less than 3 months). A single person living alone is a one-person household.*

Write down the name (or nickname) 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 an exact age unless it might be close to seven. Record the sex of the head and the sex of his/her spouse (if there is one). For each member 7-years-old or older, ask: "In the past calender-week, did [NAME] work at least one hour?", and record the response. Ask each member who worked: "In his/her main occupation, does [NAME] work as a wage/salary employee, domestic servant, or business owner who draws a salary?" Record the response. If the response is No, then ask: "In his/her main occupation, is [NAME] self-employed, a business owner who does not draw a salary, or a member of a producer's cooperative?". Record the response.

After you finish with all household members, record the exact number of members in the scorecard header next to "Number of household members". Then circle the response to the second scorecard question. Record the responses to the third and fourth scorecard questions according to the responses in the last two columns in the table below. Read aloud the remaining six questions. Always 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 7-years old, then ask: "In the past calender-week, did [NAME] work at least one hour?"**	**If [NAME] worked, then ask: "In his/her main occupation, does [NAME] work as a wage/salary employee, domestic servant, or business owner who draws a salary?"**	**If the previous answer is "No", then ask: "In his/her main occupation, is [NAME] self-employed, a business owner who does not draw a salary, or a member of a producer's cooperative?"**
1.	Head (male) Head (female)		<7 No Yes	<7 No Yes	<7 No Yes
2.	Spouse (female) Spouse (male) Other member		<7 No Yes	<7 No Yes	<7 No Yes
3.	Other		<7 No Yes	<7 No Yes	<7 No Yes
4.	Other		<7 No Yes	<7 No Yes	<7 No Yes
5.	Other		<7 No Yes	<7 No Yes	<7 No Yes
6.	Other		<7 No Yes	<7 No Yes	<7 No Yes
7.	Other		<7 No Yes	<7 No Yes	<7 No Yes
8.	Other		<7 No Yes	<7 No Yes	<7 No Yes
9.	Other		<7 No Yes	<7 No Yes	<7 No Yes
10.	Other		<7 No Yes	<7 No Yes	<7 No Yes
# HH members:	—	—	—	**Total Yes:**	**Total Yes:**

Figure 1: Conversion of scores to poverty likelihoods

	Poverty likelihood (%)														
	National lines				Intl. 2011 PPP lines				Percentile-based lines						
Score	Food	100%	150%	200%	$1.90	$3.20	$5.50	$21.70	10th	20th	40th	50th	60th	80th	90th
0–25	71.1	92.5	98.1	99.9	38.2	58.5	84.6	100.0	57.8	79.8	95.1	99.5	99.7	100.0	100.0
26–29	45.5	83.9	91.2	98.0	10.3	34.2	72.9	99.9	33.3	61.9	85.9	92.0	97.4	99.6	100.0
30–32	30.9	73.6	88.5	96.6	9.2	22.3	54.8	99.8	21.9	44.4	75.1	86.1	92.6	98.5	99.9
33–35	23.4	66.9	85.3	95.4	5.2	18.0	45.1	99.8	17.3	35.9	68.8	83.4	90.2	98.0	99.9
36–37	17.0	61.2	81.8	92.0	2.7	10.7	42.1	99.7	10.6	28.3	65.8	78.9	88.7	96.1	99.9
38–39	12.5	48.2	72.4	84.7	2.7	8.8	33.8	98.1	8.8	24.6	53.1	65.4	76.5	93.8	98.5
40–41	11.1	48.2	70.1	82.3	2.3	7.7	30.9	96.5	7.6	20.4	53.1	64.3	74.2	91.8	97.3
42–43	9.5	37.9	62.3	79.5	2.0	6.9	22.2	94.9	6.8	15.6	41.4	53.2	67.0	90.7	96.2
44–45	5.4	32.8	58.2	74.3	1.8	4.5	20.3	92.6	4.4	12.8	35.7	47.6	62.6	84.3	93.7
46–47	5.2	26.8	55.8	72.7	1.6	4.5	14.4	91.5	4.4	9.1	33.7	45.4	58.6	82.3	93.1
48–49	4.2	22.4	50.6	68.8	1.3	3.4	11.8	88.8	3.4	8.6	25.3	37.3	51.6	78.9	90.7
50–51	3.9	18.0	42.6	61.5	1.3	3.4	10.1	82.5	3.4	7.8	19.8	28.4	40.6	70.4	84.6
52–53	2.2	10.7	31.9	50.3	0.9	2.1	6.6	76.6	2.1	5.3	13.2	21.5	31.2	61.3	78.7
54–55	1.1	7.5	27.2	45.4	0.4	0.9	4.3	75.5	0.9	2.8	9.1	17.2	25.9	57.3	78.4
56–57	0.5	6.2	23.2	41.6	0.2	0.4	2.8	69.8	0.4	1.9	6.9	12.1	21.9	52.1	74.1
58–60	0.4	3.5	16.3	30.2	0.2	0.4	2.0	51.6	0.4	1.5	4.3	7.3	13.8	34.7	57.1
61–65	0.4	1.9	9.4	21.2	0.1	0.4	1.0	46.8	0.4	0.9	2.2	4.7	8.7	28.3	49.6
66–100	0.4	0.6	3.3	9.3	0.1	0.4	0.6	33.9	0.4	0.6	0.6	0.6	2.8	13.2	38.5

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 lines				Intl. 2011 PPP lines				Percentile-based lines						
	Food	100%	150%	200%	$1.90	$3.20	$5.50	$21.70	10th	20th	40th	50th	60th	80th	90th
Estimation error	+2.0	+2.4	+2.5	+0.6	–0.4	+0.3	+3.5	–2.7	+0.7	+1.8	+3.3	+1.6	–0.4	–1.8	–2.7
Margin of error	2.4	3.5	3.6	3.0	1.5	2.2	3.3	1.3	2.0	3.0	3.5	3.3	2.9	2.1	1.3
α factor	1.43	1.37	1.34	1.27	1.41	1.41	1.43	0.82	1.33	1.46	1.34	1.26	1.15	1.06	0.84

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 Bolivia

1. Introduction

The Scorocs Simple Poverty Scorecard-brand poverty-assessment tool for Bolivia 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 income surveys. A case in point is Bolivia's 2019 Household Survey (EH, *la Encuesta de Hogares*) by Bolivia's *Instituto Nacional de la Estadística* (INE). The 2019 EH has 56 pages and asks about 500 top-level questions, many of which have several follow-up questions or are repeated (for example, for each household member or for each consumer durable).

1.2 How the scorecard works

The scorecard has 10 factual questions that are drawn from the exhaustive 2019 EH. Examples include: “Is firewood, dung, manure, or llama pellets the main fuel or source of energy used for cooking?” and “Does the household possess and use a frigerator or freezer?”.

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 departments of Bolivia

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 income-based poverty in the 2019 EH.

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.

A trained 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 EH 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 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 available to 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 Income-based poverty

Bolivia's scorecard is a quantitative way to assess whether a program's participants have income below any of 15 poverty lines. The most-relevant line is Bolivia's national line (called here "100% of the national line") of about BOB28 per person per day, giving a country-wide head-count poverty rate of 37.2 percent in 2019

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] Bolivia'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 in accuracy 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 Bolivia 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 2019 EH. 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 2020−21 economic downturn due to COVID−19 changed the links between poverty and questions, but the Bolivia scorecard still uses 2019 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 our work area?"
- "Are our participants more likely to rise above a poverty line than the average poor person in our work area?"

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 2019 EH 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 estimation error for estimates of poverty rates in one time period, allowing scorecard users to adjust for it.

1.8 What is next?

Section 2: **How to convert responses to poverty likelihoods**

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

- **Poverty in a single time period**:
  - **Head-count poverty rate**
  - **Number of poor people**
- **Annual net changes in poverty across two time periods for on-going participants**:
  - **Poverty rate with one sample scored twice**
  - **Number of poor people with one sample scored twice**
  - **Poverty rate with two independent samples**
  - **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 Bolivia's 2019 EH 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**

Cited **References** appear at the end.

2. How to convert 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 a **mobile app** 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 ("BOL", pre-filled)
  - Scorecard code ("004", 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 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 department does the household live?"). If the enumerator already knows the department (as is usually the case), then the question does not need to 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 ("How many household members work in their main occupation as wage/salary employees, domestic servants, or business owners who draw a salary?")
  - The response to the fourth scorecard question ("How many household members in their main occupation are self-employed, business owners who do not draw a salary, or members of a producer's cooperative?")

- Read aloud the remaining six questions one-by-one and in order, marking the responses given by the respondent
- Do not read the response options for any scorecard question to 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

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, and 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 scorecard's header.

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

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 Bolivia's INE in the 2019 EH. 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** provides 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 study the **Interview Guide** and scrupulously follow it.

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 Bolivia's INE did in the 2019 EH.

[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 **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

			Full name	Identifier
Interview ID:	A123			
Interview date:	13JUN2022	Participant of record:	ANNA JACKSON	1V0276FZ7
Country:	BOL	Service agent:	UNKNOWN	UNKNOWN
Scorecard:	004	Service point:	EAST CLINIC	NONE
Sampling weight:	UNKNOWN	Number of household members:	FIVE	

Question	Response	Points	
1. In which department does the household live? (*record without asking*)	A. Chuquisaca	0	
	B. La Paz	4	
	C. Tarija, or Oruro	**6**	**6**
	D. Cochabamba, or Potosí	7	
	E. Pando, or Beni	11	
	F. Santa Cruz	13	
2. How many members does the household have? (*from the **Back-page** Worksheet*)	A. Six or more	0	
	B. Five	**4**	**4**
	C. Four	10	
	D. Three	16	
	E. Two	23	
	F. One	28	
3. How many household members work in their main occupation as wage/salary employees, domestic servants, or business owners who draw a salary? (*from the **Back-page Worksheet***)	**A. None**	**0**	**0**
	B. One	9	
	C. Two or more	17	
4. How many household members in their main occupation are self-employed, business owners who do not draw a salary, or members of a producer's cooperative? (*from the **Back-page Worksheet***)	A. None	0	
	B. One	**4**	**4**
	C. Two or more	12	
5. Is the female head (or spouse of the male head) covered by medical insurance?	**A. No**	**0**	**0**
	B. Yes	2	
	C. No female head (nor spouse of the male head)	4	
6. How many rooms does the household occupy, not counting bathrooms, kitchen, laundry rooms, garages, or rooms used for storage or business?	**A. One**	**0**	**0**
	B. Two or three	3	
	C. Four or more	5	
7. Is firewood, dung, manure, or llama pellets the main fuel or source of energy used for cooking?	A. Yes	0	
	B. No	**4**	**4**
8. Does the household possess and use a frigerator or freezer?	**A. No**	**0**	**0**
	B. Yes	2	
9. Does the household possess and use an old-style tube TV or one or more flat-screen TVs (plasma, LCD, LED)?	A. No	0	
	B. Only tube	**2**	**2**
	C. One flat screen (regardless of tube)	6	
	D. Two or more flat screens (regardless of tube)	9	
10. How many household members have a cell phone available for their own personal (non-business) use?	A. None	0	
	B. One	**2**	**2**
	C. Two	3	
	D. Three or more	6	

scorocs.com

Score: 6 + 4 + 0 + 4 + 0 + 0 + 4 + 0 + 2 + 2 = 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 7-years old, then ask: "In the past calender-week, did [NAME] work at least one hour?"	If [NAME] worked, then ask: "In his/her main occupation, does [NAME] work as a wage/salary employee, domestic servant, or business owner who draws a salary?"	If the previous answer is "No", then ask: "In his/her main occupation, is [NAME] self-employed, a business owner who does not draw a salary, or a member of a producer's cooperative?"
1. **ANNA**	Head (male) **Head (female)**	**38**	<7 **No** Yes	<7 No Yes	<7 No Yes
2. **BILLY**	Spouse (female) Spouse (male) **Other member**	**24**	<7 No **Yes**	<7 **No** Yes	<7 No **Yes**
3. **CHARLES**	**Other**	**18**	<7 **No** Yes	<7 No Yes	<7 No Yes
4. **DARLA**	**Other**	**16**	<7 **No** Yes	<7 No Yes	<7 No Yes
5. **EUGENE**	**Other**	**4**	**<7** No Yes	<7 No Yes	<7 No Yes
6.	Other		<7 No Yes	<7 No Yes	<7 No Yes
7.	Other		<7 No Yes	<7 No Yes	<7 No Yes
8.	Other		<7 No Yes	<7 No Yes	<7 No Yes
9.	Other		<7 No Yes	<7 No Yes	<7 No Yes
10.	Other		<7 No Yes	<7 No Yes	<7 No Yes
# HH members: **FIVE**	—	—	—	**Total Yes: None**	**Total Yes: ONE**

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 all supported poverty lines, **Figure 1** lists poverty likelihoods by score range. A score of 22 falls in the first range of 0−25. For 100% of the national poverty line, the poverty likelihood for scores of 0−25 is 92.5 percent. That is, the scorecard estimates that 92.5 percent of households in Bolivia with a score of 0−25 have income below 100% of the national line.

Figure 5: The first example household's score of 22 corresponds with a poverty likelihood of 92.5 percent for 100% of the national line (excerpted from Figure 1)

	Poverty likelihood (%)			
		National lines		
Score	Food	100%	150%	200%
0−25	71.1	92.5	98.1	99.9
26−29	45.5	83.9	91.2	98.0
30−32	30.9	73.6	88.5	96.6
33−35	23.4	66.9	85.3	95.4
36−37	17.0	61.2	81.8	92.0
38−39	12.5	48.2	72.4	84.7
40−41	11.1	48.2	70.1	82.3
42−43	9.5	37.9	62.3	79.5
44−45	5.4	32.8	58.2	74.3
. . .	. . .	. . .	. . .	. . .

Figure 6: Second example household, filled-in scorecard

			Full name	Identifier
Interview ID:	B456			
Interview date:	30JUN2022	Participant of record:	JOHN BROWN	2W3120ZG8
Country:	BOL	Service agent:	UNKNOWN	UNKNOWN
Scorecard:	004	Service point:	EAST CLINIC	NONE
Sampling weight:	UNKNOWN	Number of household members:	SEVEN	

Question	Response	Points	
1. In which department does the household live? (*record without asking*)	A. Chuquisaca	0	**0**
	B. La Paz	4	
	C. Tarija, or Oruro	6	
	D. Cochabamba, or Potosí	7	
	E. Pando, or Beni	11	
	F. Santa Cruz	13	
2. How many members does the household have? (*from the **Back-page** Worksheet*)	A. Six or more	0	**0**
	B. Five	4	
	C. Four	10	
	D. Three	16	
	E. Two	23	
	F. One	28	
3. How many household members work in their main occupation as wage/salary employees, domestic servants, or business owners who draw a salary? (*from the **Back-page Worksheet***)	A. None	0	
	B. One	9	**9**
	C. Two or more	17	
4. How many household members in their main occupation are self-employed, business owners who do not draw a salary, or members of a producer's cooperative? (*from the **Back-page Worksheet***)	A. None	0	
	B. One	4	**4**
	C. Two or more	12	
5. Is the female head (or spouse of the male head) covered by medical insurance?	A. No	0	
	B. Yes	2	**2**
	C. No female head (nor spouse of the male head)	4	
6. How many rooms does the household occupy, not counting bathrooms, kitchen, laundry rooms, garages, or rooms used for storage or business?	A. One	0	
	B. Two or three	3	**3**
	C. Four or more	5	
7. Is firewood, dung, manure, or llama pellets the main fuel or source of energy used for cooking?	A. Yes	0	
	B. No	4	**4**
8. Does the household possess and use a frigerator or freezer?	A. No	0	
	B. Yes	2	**2**
9. Does the household possess and use an old-style tube TV or one or more flat-screen TVs (plasma, LCD, LED)?	A. No	0	
	B. Only tube	2	
	C. One flat screen (regardless of tube)	6	**6**
	D. Two or more flat screens (regardless of tube)	9	
10. How many household members have a cell phone available for their own personal (non-business) use?	A. None	0	
	B. One	2	**2**
	C. Two	3	
	D. Three or more	6	

scorocs.com

Score: 0 + 0 + 9 + 4 + 2 + 3 + 4 + 2 + 6 + 2 = 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 7-years old, then ask: "In the past calender-week, did [NAME] work at least one hour?"	If [NAME] worked, then ask: "In his/her main occupation, does [NAME] work as a wage/salary employee, domestic servant, or business owner who draws a salary?"	If the previous answer is "No", then ask: "In his/her main occupation, is [NAME] self-employed, a business owner who does not draw a salary, or a member of a producer's cooperative?"
1. **ALBERT**	Head (male) Head (female)	**35**	<7 No **Yes**	<7 No **Yes**	<7 No Yes
2. **BERNITA**	Spouse (female) Spouse (male) Other member	**34**	<7 No **Yes**	<7 **No** Yes	<7 No **Yes**
3. **CARLOS**	Other	**16**	<7 No **Yes**	<7 **No** Yes	<7 **No** Yes
4. **DARLENE**	Other	**15**	<7 **No** Yes	<7 No Yes	<7 No Yes
5. **EVELYN**	Other	**11**	<7 **No** Yes	<7 No Yes	<7 No Yes
6. **FRANCINE**	Other	**6**	**<7** No Yes	<7 No Yes	<7 No Yes
7. **GEORGE**	Other	**3**	**<7** No Yes	<7 No Yes	<7 No Yes
8.	Other		<7 No Yes	<7 No Yes	<7 No Yes
9.	Other		<7 No Yes	<7 No Yes	<7 No Yes
10.	Other		<7 No Yes	<7 No Yes	<7 No Yes
# HH members: **SEVEN**	—	—	—	**Total Yes: ONE**	**Total Yes: ONE**

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**).

For all supported poverty lines, **Figure 1** lists poverty likelihoods by score range. A score of 32 falls in the third range of 30−32. For 100% of the national poverty line, the poverty likelihood for scores of 30−32 is 73.6 percent. That is, the scorecard estimates that 73.6 percent of households in Bolivia with a score of 30−32 have income below 100% of the national line.

Figure 8: The second example household's score of 32 corresponds with a poverty likelihood of 73.6 percent for 100% of the national line (excerpted from Figure 1)

	Poverty likelihood (%)			
		National lines		
Score	Food	100%	150%	200%
0−25	71.1	92.5	98.1	99.9
26−29	45.5	83.9	91.2	98.0
30−32	30.9	73.6	88.5	96.6
33−35	23.4	66.9	85.3	95.4
36−37	17.0	61.2	81.8	92.0
38−39	12.5	48.2	72.4	84.7
40−41	11.1	48.2	70.1	82.3
42−43	9.5	37.9	62.3	79.5
44−45	5.4	32.8	58.2	74.3
. . .	. . .	. . .	. . .	. . .

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 Poverty in a single time period

3.1.1 Head-count poverty rate

The *head-count poverty rate* is the share of people in participating households in which total household income (divided by the number of members 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 in a pro-poor program that operates throughout Bolivia enrolls 1,000 in-coming households in calendar-year 2022, 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 +2.4 percentage points (**Figure 2**).

[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 first example household has five members and is interviewed on June 13, 2022 (**Figure 3** and **Figure 4**). Its score of 22 corresponds with a poverty likelihood of 92.5 percent.

The second example household has seven members and is interviewed on June 30, 2022 (**Figure 6** and **Figure 7**). Its score of 32 corresponds with a poverty likelihood of 73.6 percent.

The estimated head-count poverty rate for the population of in-coming households in the 2022 calendar-year cohort 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:

$$\frac{5 \cdot 0.925 + 7 \cdot 0.736}{5 + 7} - (+0.024) \approx \frac{9.78}{12} - 0.024 \approx 0.791 = 79.1 \text{ percent.}$$

The five in the "$5 \cdot 0.925$" term is the number of members (household size) in the first household, and 0.925 is the first household's estimated poverty likelihood as a proportion.

In the same way, the seven in "$7 \cdot 0.736$" is the number of members in the second household, and 0.736 is the second household's estimated poverty likelihood.

The "5 + 7" is the sum of the weights—that is, the number of household members—across the two sampled households.

The "+0.024" is the scorecard's estimation error for this poverty line (**Figure 2**). Because unadjusted estimates tend to be too high by 2.4 percentage points, they are adjusted downwards by subtracting +2.4. This is akin to how an archer whose arrows tend to miss a little to the right of the bulls-eye will adjust his/her aim to be a little to the left of the bulls-eye.

The estimated head-count poverty rate for the population is 71.9 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 are done with the **ProveIt™-brand reporting and analysis tool** 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 $(5 \cdot 22 + 7 \cdot 32) \div (5 + 7) \approx 28$. This average score of 28 corresponds to a poverty likelihood for 100% of the national poverty line of 83.9 percent (**Figure 1**), giving an error-adjusted poverty rate of 83.9 − (+2.4) = 81.5 percent. This differs from the 79.1 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 participant of record**	**Number of household members**	**Score**	**Poverty likelihood (%)**	**Estimated number of poor household members**
2	Baseline	13-Jun-22	1V0276FZ7	5	22	92.5	4.63 = (D2*F2)/100
3	Baseline	30-Jun-22	2W3120ZG8	7	32	73.6	5.15 = (D3*F3)/100
4			**Sum:**	12 = SUM(D2:D3)			9.78 = SUM(G2:G3)
5			**Average:**	6.0 = AVERAGE(D2:D3)			
6							
7	**Estimated scorecard error for this poverty line (percentage points):**						+2.4
8							
9				**Estimated head-count poverty rate (%):**			79.1 = (G4/D4)*100-G7
10							
11				**Households in the population:**			1,000
12							
13				**People in households in the population:**			6,000 = G11*D5
14							
15				**Number of poor people in population:**			4,745 = (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 program's work area.

To help with benchmarking poverty-rate estimates, **Figure 10** reports head-count poverty rates from the 2019 EH for all 15 supported poverty lines by urban/rural/all for Bolivia overall and for each of its nine departments. For Bolivia overall, the head-count poverty rate for 100% of the national line is 37.2 percent. Thus, the example program is pro-poor in the sense that its in-coming participants have an above-average estimated poverty rate (79.1 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 areas in **Figure 10** begin with Bolivia overall, followed by the nine departments in alphabetical order.

[14] The Bolivia scorecard uses a definition of *poverty* based on income. 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: (Bolivia overall, Beni, and Chuquisaca): Poverty lines and head-count poverty rates by urban/rural/all in 2019

	Dept./ Area	Line or Rate	*n*	Poverty lines (BOB) and poverty rates (%)														
				National lines				Intl. 2011 PPP lines				Percentile-based lines						
				Food	100%	150%	200%	$1.90	$3.20	$5.50	$21.70	10th	20th	40th	50th	60th	80th	90th
Bolivia	Urban	Line	9,059	15	30	50	66	8	14	24	93	14	20	32	38	46	71	100
		Rate		6.4	31.4	58.0	73.0	1.7	5.4	19.9	87.4	5.3	14.2	34.2	44.8	55.4	77.9	89.2
	Rural	Line	2,780	13	22	26	35	6	10	17	68	10	15	23	28	34	52	73
		Rate		27.8	50.8	59.7	71.8	10.7	21.3	39.2	90.9	20.9	33.5	53.2	61.9	70.7	84.8	91.8
	All	Line	11,839	14	28	43	57	8	13	22	86	12	19	29	35	42	65	92
		Rate		12.9	37.2	58.5	72.6	4.4	10.2	25.7	88.5	10.0	20.0	40.0	50.0	60.0	80.0	90.0
Beni	Urban	Line	507	13	27	39	52	7	12	21	83	12	18	28	34	41	63	89
		Rate		3.4	30.6	50.4	68.3	1.1	3.2	15.3	88.5	3.2	10.3	34.8	44.0	53.6	76.1	90.0
	Rural	Line	156	13	22	26	35	6	10	17	68	10	15	23	28	34	52	73
		Rate		31.0	55.7	64.6	73.6	15.2	26.5	42.0	91.3	26.5	37.0	58.8	65.1	70.5	83.4	92.0
	All	Line	663	13	26	36	48	7	12	20	79	12	17	27	32	39	60	85
		Rate		10.4	36.9	54.0	69.7	4.6	9.1	22.0	89.2	9.1	17.0	40.8	49.3	57.9	77.9	90.5
Chuquisaca	Urban	Line	529	16	33	58	77	9	15	26	101	15	22	34	41	50	77	108
		Rate		9.9	43.4	72.6	81.8	1.9	8.9	30.4	91.9	7.9	21.2	45.4	57.4	66.7	83.5	93.3
	Rural	Line	395	13	22	26	35	6	10	17	68	10	15	23	28	34	52	73
		Rate		39.9	65.6	73.0	76.2	20.6	33.3	52.3	93.5	32.3	45.1	67.4	73.7	75.0	87.8	93.6
	All	Line	924	14	27	42	57	7	13	22	85	12	18	29	35	42	65	91
		Rate		24.6	54.3	72.8	79.1	11.1	20.9	41.1	92.7	19.9	32.9	56.2	65.4	70.8	85.6	93.4

Source: 2019 EH

All poverty lines are BOB per-person and per-day in prices in Bolivia as a whole on average in during the 2019 EH fieldwork.

Figure 10: (Cochabamba, Oruro, and Pando): Poverty lines and head-count poverty rates by urban/rural/all in 2019

Dept./ Area	Line or Rate	*n*	Poverty lines (BOB) and poverty rates (%)														
			National lines				Intl. 2011 PPP lines				Percentile-based lines						
			Food	100%	150%	200%	$1.90	$3.20	$5.50	$21.70	10th	20th	40th	50th	60th	80th	90th
Cochabamba	**Urban**	Line	16	32	58	77	9	15	26	101	15	22	34	41	50	77	108
		Rate	7.4	35.3	66.5	79.4	1.9	6.0	22.9	89.8	6.0	14.9	38.2	48.6	59.4	81.1	91.5
		n	1,843														
	Rural	Line	13	22	26	35	6	10	17	68	10	15	23	28	34	52	73
		Rate	27.5	54.7	62.9	77.6	8.3	22.6	38.6	95.5	21.5	33.4	55.8	66.3	76.5	90.9	95.6
		n	348														
	All	Line	15	29	48	64	8	13	23	91	13	20	31	37	45	69	97
		Rate	13.4	41.1	65.4	78.8	3.8	10.9	27.6	91.5	10.6	20.4	43.5	53.9	64.5	84.1	92.7
		n	2,191														
Oruro	**Urban**	Line	14	29	45	60	8	13	23	89	13	19	30	36	44	68	95
		Rate	5.5	24.7	50.6	68.3	1.7	4.8	18.8	87.4	4.8	12.9	27.6	38.7	49.6	75.2	88.9
		n	516														
	Rural	Line	13	22	26	35	6	10	17	68	10	15	23	28	34	52	73
		Rate	24.1	55.9	61.9	76.4	9.2	18.0	44.1	91.0	17.8	33.8	58.7	64.6	75.9	88.0	91.9
		n	360														
	All	Line	14	26	39	51	7	12	21	82	12	18	28	34	41	63	88
		Rate	11.8	35.3	54.5	71.1	4.3	9.3	27.4	88.6	9.2	20.0	38.2	47.5	58.6	79.6	89.9
		n	876														
Pando	**Urban**	Line	14	30	49	66	8	14	24	93	14	20	32	38	46	71	100
		Rate	2.8	27.6	49.6	69.3	0.6	2.3	19.0	84.0	2.3	13.8	29.7	35.4	44.9	75.8	87.5
		n	320														
	Rural	Line	13	22	26	35	6	10	17	68	10	15	23	28	34	52	73
		Rate	16.8	35.9	42.2	58.9	3.9	8.8	24.7	89.2	7.0	20.7	37.4	45.9	58.1	80.0	89.6
		n	216														
	All	Line	14	27	40	53	7	12	21	83	12	18	28	34	41	63	89
		Rate	8.6	31.0	46.6	65.0	2.0	5.0	21.3	86.2	4.3	16.6	32.9	39.7	50.3	77.5	88.3
		n	536														

Source: 2019 EH

All poverty lines are BOB per-person and per-day in prices in Bolivia as a whole on average in during the 2019 EH fieldwork.

Figure 10: (La Paz, Potosí, and Santa Cruz): Poverty lines and head-count poverty rates by urban/rural/all in 2019

Dept./ Area	Line or Rate	*n*	Poverty lines (BOB) and poverty rates (%)														
			National lines				Intl. 2011 PPP lines				Percentile-based lines						
			Food	100%	150%	200%	$1.90	$3.20	$5.50	$21.70	10th	20th	40th	50th	60th	80th	90th
La Paz	Urban	Line	16	33	61	82	9	15	26	104	15	22	35	42	51	79	111
		Rate	8.3	37.1	71.2	84.0	2.0	6.8	23.7	92.0	6.7	16.8	40.1	50.9	62.5	84.0	93.3
		n	2,520														
	Rural	Line	13	22	26	35	6	10	17	68	10	15	23	28	34	52	73
		Rate	32.5	55.2	63.7	73.1	13.1	23.3	44.9	91.5	23.0	39.2	58.3	66.1	73.0	86.7	92.5
		n	360														
	All	Line	15	30	50	66	8	14	23	92	13	20	31	38	46	70	99
		Rate	16.1	43.0	68.8	80.5	5.6	12.2	30.6	91.9	12.0	24.1	46.0	55.8	65.9	84.9	93.0
		n	2,880														
Potosí	Urban	Line	14	29	46	61	8	13	23	90	13	19	31	37	44	68	96
		Rate	8.0	37.8	63.0	77.4	1.4	4.8	29.7	88.6	4.8	21.4	40.7	53.8	61.7	82.9	89.5
		n	348														
	Rural	Line	13	22	26	35	6	10	17	68	10	15	23	28	34	52	73
		Rate	30.9	51.3	59.8	71.4	11.1	24.6	41.5	88.5	24.6	35.3	53.1	61.2	70.2	81.2	89.4
		n	408														
	All	Line	13	25	35	46	7	11	20	78	11	17	26	32	38	59	83
		Rate	21.1	45.5	61.2	74.0	7.0	16.1	36.4	88.6	16.1	29.4	47.8	58.0	66.6	81.9	89.4
		n	756														
Santa Cruz	Urban	Line	13	27	39	52	7	12	21	83	12	18	28	34	41	63	89
		Rate	4.5	23.7	43.8	61.3	1.5	4.1	13.5	82.1	4.0	10.3	26.3	36.6	47.1	71.0	84.5
		n	1,747														
	Rural	Line	13	22	26	35	6	10	17	68	10	15	23	28	34	52	73
		Rate	14.0	31.0	43.2	61.5	4.8	9.1	23.1	85.0	9.1	19.0	34.6	45.5	59.0	77.6	86.9
		n	263														
	All	Line	13	26	37	49	7	12	20	80	12	17	27	33	40	61	86
		Rate	6.1	24.9	43.7	61.3	2.0	5.0	15.1	82.6	4.8	11.7	27.6	38.1	49.1	72.2	84.9
		n	2,010														

Source: 2019 EH

All poverty lines are BOB per-person and per-day in prices in Bolivia as a whole on average in during the 2019 EH fieldwork.

Figure 10: (Tarija): Poverty lines and head-count poverty rates by urban/rural/all in 2019

	Dept./ Area	Line or Rate	*n*	Poverty lines (BOB) and poverty rates (%)														
				National lines				Intl. 2011 PPP lines				Percentile-based lines						
				Food	100%	150%	200%	$1.90	$3.20	$5.50	$21.70	10th	20th	40th	50th	60th	80th	90th
Tarija	Urban	Line	729	15	31	52	69	8	14	24	95	14	20	32	39	47	72	101
		Rate		7.8	34.6	59.2	74.5	1.4	6.2	23.0	87.6	6.2	17.3	37.9	48.5	56.7	79.4	89.9
	Rural	Line	274	13	22	26	35	6	10	17	68	10	15	23	28	34	52	73
		Rate		22.4	45.0	56.9	69.8	6.3	17.8	29.3	93.7	17.8	25.2	47.5	58.5	68.8	80.4	94.6
	All	Line	1,003	14	28	43	58	8	13	22	86	13	19	29	35	43	66	92
		Rate		12.5	38.0	58.4	73.0	3.0	9.9	25.1	89.5	9.9	19.9	41.0	51.8	60.6	79.7	91.4

Source: 2019 EH

All poverty lines are BOB per-person and per-day in prices in Bolivia as a whole on average in during the 2019 EH fieldwork.

3.1.2 Number of poor people

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 the two-household example with simple random sampling, this is the equal-weighted average of the number of people in the sampled households:

$$\frac{5+7}{1+1} = \frac{12}{2} = 6.0 \text{ people.}$$

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, with an estimated average of 6.0 members per household. The estimated number of people in the households of in-coming participants is then 1,000 · 6.0 = 6,000.

The third and final step is to multiply the estimated poverty rate (here, 79.1 percent, or 0.791) by the estimated number of people in in-coming households (here, 6,000). This gives 6,000 · 0.791 ≈ 4,745 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 estimates from a scorecard.

3.2 Annual net changes in poverty 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.

3.2.1 Poverty rate 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 six members (rather than five as at baseline) and is scored a second time on August 13, 2025, which is 1,157 days (about 3.17 years) after its first interview on June 13, 2022. Its score is now 27 (rather than 22), so its poverty likelihood for 100% of the national line is 83.9 percent (**Figure 1**).

Suppose that the second household now has eight members (rather than seven as at baseline) and is scored a second time on May 15, 2025, which is 1,050 days (about 2.88 years) after its first interview on June 30, 2022. Its score is now 35 (rather than 32), so its poverty likelihood has decreased from 73.6 to 66.9 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. The remaining 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 −7.5 percentage points:

$$\frac{\left(\frac{5+6}{2}\right)\cdot(0.839-0.925)+\left(\frac{7+8}{2}\right)\cdot(0.669-0.736)}{\left(\frac{5+6}{2}\right)+\left(\frac{7+8}{2}\right)}\approx\frac{-0.473+-0.502}{13.0}\approx-0.075.$$

The estimated head-count poverty rate decreased (improved) by 7.5 *percentage points* (not by 7.5 *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:

$$\frac{\left(\frac{5+6}{2}\right)\cdot 3.17+\left(\frac{7+8}{2}\right)\cdot 2.88}{\left(\frac{8+7}{2}\right)+\left(\frac{5+6}{2}\right)}\approx\frac{17.44+21.60}{13}\approx 3.00 \text{ years}.$$

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:
−7.5 ÷ 3.00 ≈ −2.5 percentage points per year.[19] The negative change means that poverty decreased (improved).[20]

In practice, the calculations are done with the **ProveIt™-brand reporting and analysis tool** or a spreadsheet modelled on **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 percent, and the rate after two years is
0.40 − 0.10 = 0.30 = 30 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	Interview date			Number of household members			Member-years	Score		Poverty likelihood (%)		Estimated net change in
2	participant of record	Baseline	Follow-up	Years between interviews	Baseline	Follow-up	Average:	between interviews	Baseline	Follow-up	Baseline	Follow-up	number of poor household members
3	1V0276FZ7	13-Jun-2022	13-Aug-2025	3.17 = (C3-B3)/365	5	6	5.50 = (E3+F3)/2	17.43 = D3*G3	22	27	92.5	83.9	−0.473 = G3*(L3-K3)/100
4	2W3120ZG8	30-Jun-2022	15-May-2025	2.88 = (C4-B4)/365	7	8	7.50 = (E4+F4)/2	21.58 = D4*G4	32	35	73.6	66.9	−0.502 = G4*(L4-K4)/100
5				**Average:**	6.0 = AVERAGE(E3:E4)	7.0 = AVERAGE(F3:F4)	**Sum:**	39.01 = SUM(H3:H4)					−0.975 = SUM(M3:M4)
6													
7						**Estimated net change in head-count poverty rate (percentage points), follow-up versus baseline:**								−7.5 = M5/(E5+F5)*100
8													
9								**Household-size-weighted average years between interviews:**					3.00 = H6/(E5+F5)
10													
11							**Estimated annual net change in head-count poverty rate (percentage points):**						−2.5 = 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:**					5,450 = (E5*M13+F5*M14)/2
17													
18								**Estimated annual net change in the number of poor people:**					−136 = M16*M11/100
19	Rows of data are sorted by the ID of the participant of record.												

3.2.2 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 2022 of in-coming households in the calendar-year 2022 cohort was 1,000. By the end of the follow-up period of calendar-year 2025, 300 had dropped out, leaving 700 from the 2022 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 households interviewed at baseline and follow-up. In a given round,

---

[21] This assumption rarely holds. On the one hand, the households that benefit most from a 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 change 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 change. 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—commonly become its *de facto* beneficiaries (**Schreiner**, 1997).

the number of participating people is the average household size for that round's interviewed households (in the example, 6.0 at baseline and 7.0 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). In the example, this is $\frac{6.0 \cdot 1{,}000 + 7.0 \cdot 700}{1+1} = 5{,}450$ people.

The second and last step is to multiply the estimated annual change in the poverty rate (here, about −2.5 percentage points, or −0.025) by the estimated average number of on-going participants (here, 5,450). This gives an estimate of the annual net change in the number of poor people by 100% of the national line of $-0.025 \cdot 5{,}450 \approx -136$ people.[22] This negative change is a decrease (improvement) in poverty; there are about 136 fewer poor people in participating households in this cohort each year.

3.2.3 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 overall or for its work area (such as those in **Figure 10**).

---

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

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]

3.2.4 Poverty rate 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, 2025. It has four members, a score of 26, and a poverty likelihood by 100% of the national line of 83.9 percent (**Figure 1**).

The fourth household is interviewed on April 4, 2025. It has three members, a score of 36, and a poverty likelihood of 61.2 percent.

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:

$$\frac{4 \cdot 0.839 + 3 \cdot 0.612}{4+3} \approx \frac{3.36 + 1.84}{7} \approx 0.742 = 74.2 \text{ percent.}$$

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

The estimated annual net change in the head-count poverty rate of on-going participants is then the difference between the (unadjusted) poverty-rate estimates at follow-up (74.2 percent) versus at baseline (81.5 percent),[25] divided by the difference (in years) between the household-size-weighted average of follow-up interview dates (March 16, 2025) versus the household-size-weighted average of baseline interview dates (June 22, 022). These two average dates differ by about 998 days or about 2.73 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 about (74.2 − 81.5) ÷ 2.73 ≈ −2.7 percentage points per year.

In practice, the calculations are done with the **ProveIt™-brand reporting and analysis tool** or a spreadsheet modelled on **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 81.5 percent, not 81.5 − (+2.4) = 79.1 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 participant of record	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-2022	5	05-Apr-2512 = C2*D2	22	92.5	4.63 = D2*G2/100
3	Baseline	2W3120ZG8	30-Jun-2022	7	29-Jun-2757 = C2*D2	32	73.6	5.15 = D3*G3/100
4	Follow-up	3XA76T21L	3-Mar-2025	4	10-Sep-2400 = C2*D2	26	83.9	3.36 = D4*G4/100
5	Follow-up	4Y8Y3EQS9	4-Apr-2025	3	13-Oct-2275 = C2*D2	36	61.2	1.84 = D5*G5/100
6		**Sum baseline:**		12 = SUM(D2:D3)				9.78 = SUM(H2:H3)
7		**Sum follow-up:**		7 = SUM(D4:D5)				5.19 = SUM(H4:H5)
8		**Average baseline:**		6.0 = AVERAGE(D2:D3)	22-Jun-2022 = SUM(E2:E3)/D6			
9		**Average follow-up:**		3.5 = AVERAGE(D4:D5)	16-Mar-2025 = SUM(E4:E5)/D7			
10								
11					**Estimated baseline poverty rate (%):**			81.5 = H6/D6*100
12					**Estimated follow-up poverty rate (%):**			74.2 = H7/D7*100
13								
14				**Average years between follow-up and baseline interviews:**				2.73 = (E9-E8)/365
15								
16			**Estimated annual net change in head-count poverty rate (percentage points):**					–2.7 = (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:**				4,225 = (D8*H18+D9*H19)/2
22								
23				**Estimated annual net change in the number of poor people:**				–113 = H21*H16/100
24	Rows of data are sorted by Survey, then by Interview date, then by the ID of the participant of record.							

3.2.5 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 2022 cohort in 2022 is 1,000 in-coming households. By the end of the 2025 follow-up period, 300 households dropped out, leaving 700 from the 2022 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, 6.0 at baseline and 3.5 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

$$\frac{6.0 \cdot 1{,}000 + 3.5 \cdot 700}{1+1} = 4{,}225 \text{ people.}$$

The second and last step is to multiply the estimated annual net change in the head-count poverty rate (here, −2.7 percentage points, or −0.027) by the estimated number of on-going participants (here, 4,225). For 100% of the national line, this gives an annual net change in the number of poor people of about $-0.027 \cdot 4{,}225 \approx -113$ people per year. This negative change is a (non-compounded) decrease in poverty; the number of poor people in participating households decreases (improves) by about 113 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. Still, 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 do 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
- Third-party 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 Bolivia's INE in the 2019 EH, 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 some 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 with a **web app** and then uploaded to a database[28]

4.4 How to calculate estimates and report/analyze them

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

[27] **Schreiner**, 2015b.

[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 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 interview households

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 multiple time 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 smaller margins of error. In addition, this approach may 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 program-selected cut-off are labeled *targeted* and given one type of treatment. People in households with scores above the 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 income below a poverty line). Poverty status is a fact that is defined by whether income 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 income below an externally-defined poverty line). Examples of such labels 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 income 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 income 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

Observed poverty status	Targeting segment: Targeted	Targeting segment: Not targeted
Poor	**Inclusion** Poor correctly targeted	**Undercoverage** Poor mistakenly not targeted
Non-poor	**Leakage** Non-poor mistakenly targeted	**Exclusion** Non-poor correctly not targeted

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 Bolivia:

- 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 lines				Intl. 2011 PPP lines				Percentile-based lines						
		Food	100%	150%	200%	$1.90	$3.20	$5.50	$21.70	10th	20th	40th	50th	60th	80th	90th
<=25	4.8	3.3	4.4	4.8	4.8	1.9	3.2	4.2	4.8	3.2	3.9	4.5	4.8	4.8	4.8	4.8
<=29	9.7	5.0	7.9	8.7	9.5	2.6	4.6	7.0	9.7	4.6	6.1	8.0	8.8	9.4	9.6	9.7
<=32	14.8	7.1	11.9	13.3	14.4	3.2	6.4	10.1	14.8	6.3	9.0	12.2	13.2	14.0	14.6	14.8
<=35	20.9	8.2	15.7	18.2	20.0	3.5	7.1	12.9	20.8	7.0	11.0	16.1	18.1	19.3	20.4	20.8
<=37	25.1	9.3	18.5	21.8	23.7	3.6	8.0	15.1	25.0	7.7	12.8	19.1	21.5	22.8	24.6	25.0
<=39	31.0	10.2	21.7	26.6	29.2	3.8	8.7	17.5	30.7	8.5	14.6	22.5	25.8	27.9	30.1	30.8
<=41	37.3	10.8	25.1	31.4	34.7	4.0	9.2	19.4	37.0	8.9	16.0	26.1	30.2	33.0	35.9	37.0
<=43	44.2	11.4	27.7	36.0	40.4	4.1	9.6	20.7	43.6	9.3	16.9	28.9	34.0	37.9	42.5	43.7
<=45	51.6	11.8	30.5	41.0	46.4	4.2	9.9	22.2	50.7	9.6	17.8	31.9	38.2	42.7	49.2	51.0
<=47	59.2	12.0	32.5	45.5	52.1	4.3	10.0	23.3	57.7	9.7	18.3	34.2	41.7	47.5	55.7	58.0
<=49	65.8	12.1	34.0	49.1	56.9	4.4	10.1	23.9	63.6	9.8	18.6	36.0	44.2	50.8	61.0	64.0
<=51	71.8	12.2	35.0	51.8	60.7	4.4	10.2	24.3	68.7	9.9	18.9	37.2	46.1	53.6	65.4	69.3
<=53	78.3	12.3	36.2	54.7	64.7	4.4	10.2	24.8	74.3	9.9	19.2	38.5	48.0	56.4	70.0	75.0
<=55	82.6	12.4	36.4	56.0	67.1	4.5	10.3	25.0	77.8	10.0	19.3	38.8	48.5	57.7	72.7	78.6
<=57	87.6	12.4	36.8	57.1	69.3	4.5	10.3	25.2	81.6	10.0	19.4	39.3	49.3	58.8	75.5	82.5
<=60	93.1	12.4	36.9	58.0	71.2	4.5	10.3	25.2	85.0	10.0	19.4	39.4	49.6	59.7	77.9	86.3
<=65	97.4	12.4	37.0	58.4	72.4	4.5	10.3	25.3	87.5	10.0	19.5	39.5	49.7	60.1	79.4	89.0
<=100	100.0	12.4	37.0	58.5	72.6	4.5	10.3	25.3	88.3	10.1	19.5	39.5	49.8	60.1	79.9	89.9

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 lines				Intl. 2011 PPP lines				Percentile-based lines						
		Food	100%	150%	200%	$1.90	$3.20	$5.50	$21.70	10th	20th	40th	50th	60th	80th	90th
<=25	4.8	9.2	32.6	53.7	67.8	2.6	7.1	21.1	83.5	6.8	15.6	35.0	45.0	55.3	75.1	85.1
<=29	9.7	7.5	29.1	49.8	63.1	1.9	5.7	18.3	78.6	5.4	13.4	31.5	41.0	50.8	70.3	80.2
<=32	14.8	5.4	25.1	45.2	58.3	1.3	4.0	15.2	73.5	3.8	10.5	27.3	36.5	46.1	65.3	75.1
<=35	20.9	4.3	21.3	40.3	52.7	1.0	3.2	12.5	67.5	3.1	8.4	23.4	31.7	40.8	59.4	69.1
<=37	25.1	3.2	18.5	36.7	48.9	0.9	2.4	10.2	63.3	2.4	6.7	20.4	28.3	37.3	55.3	64.9
<=39	31.0	2.2	15.3	31.9	43.4	0.7	1.6	7.8	57.6	1.6	4.9	17.0	23.9	32.2	49.8	59.1
<=41	37.3	1.6	11.9	27.1	37.9	0.5	1.2	6.0	51.3	1.1	3.5	13.4	19.6	27.2	43.9	52.9
<=43	44.2	1.0	9.3	22.5	32.2	0.4	0.7	4.6	44.7	0.7	2.6	10.6	15.7	22.2	37.4	46.2
<=45	51.6	0.7	6.6	17.5	26.2	0.3	0.5	3.1	37.6	0.5	1.7	7.6	11.5	17.4	30.7	38.9
<=47	59.2	0.4	4.5	13.0	20.5	0.2	0.3	2.0	30.6	0.3	1.1	5.3	8.1	12.7	24.2	31.9
<=49	65.8	0.3	3.1	9.4	15.8	0.1	0.3	1.4	24.7	0.3	0.9	3.5	5.6	9.3	18.9	25.9
<=51	71.8	0.2	2.0	6.7	11.9	0.1	0.2	1.0	19.6	0.2	0.6	2.3	3.7	6.6	14.5	20.6
<=53	78.3	0.1	0.9	3.8	7.9	0.0	0.1	0.5	14.0	0.1	0.3	1.0	1.8	3.8	9.9	14.9
<=55	82.6	0.1	0.6	2.5	5.5	0.0	0.0	0.3	10.5	0.0	0.2	0.7	1.2	2.5	7.2	11.3
<=57	87.6	0.0	0.2	1.3	3.3	0.0	0.0	0.2	6.7	0.0	0.1	0.2	0.5	1.3	4.4	7.4
<=60	93.1	0.0	0.1	0.5	1.4	0.0	0.0	0.1	3.3	0.0	0.1	0.1	0.1	0.5	2.0	3.6
<=65	97.4	0.0	0.0	0.1	0.3	0.0	0.0	0.0	0.8	0.0	0.0	0.0	0.0	0.0	0.5	0.9
<=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

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 lines				Intl. 2011 PPP lines				Percentile-based lines						
		Food	100%	150%	200%	$1.90	$3.20	$5.50	$21.70	10th	20th	40th	50th	60th	80th	90th
<=25	4.8	1.5	0.4	0.1	0.0	2.9	1.6	0.6	0.0	1.6	1.0	0.3	0.0	0.0	0.0	0.0
<=29	9.7	4.8	1.8	1.0	0.2	7.1	5.1	2.8	0.0	5.1	3.6	1.7	1.0	0.4	0.2	0.0
<=32	14.8	7.7	2.9	1.5	0.4	11.5	8.4	4.7	0.0	8.5	5.8	2.6	1.5	0.7	0.2	0.0
<=35	20.9	12.7	5.2	2.7	0.9	17.4	13.8	8.0	0.1	13.9	9.8	4.8	2.8	1.6	0.4	0.1
<=37	25.1	15.9	6.6	3.4	1.4	21.5	17.2	10.0	0.1	17.5	12.4	6.0	3.6	2.3	0.6	0.1
<=39	31.0	20.7	9.2	4.4	1.8	27.2	22.2	13.5	0.2	22.5	16.4	8.5	5.1	3.0	0.9	0.2
<=41	37.3	26.4	12.2	5.9	2.6	33.3	28.1	17.9	0.3	28.4	21.3	11.2	7.1	4.3	1.4	0.3
<=43	44.2	32.7	16.5	8.2	3.8	40.0	34.5	23.4	0.6	34.8	27.3	15.2	10.1	6.3	1.7	0.5
<=45	51.6	39.8	21.2	10.6	5.2	47.4	41.7	29.4	0.9	42.0	33.8	19.7	13.4	8.9	2.4	0.7
<=47	59.2	47.1	26.6	13.7	7.0	54.8	49.1	35.8	1.4	49.4	40.8	25.0	17.5	11.7	3.5	1.1
<=49	65.8	53.7	31.9	16.8	9.0	61.5	55.8	41.9	2.2	56.0	47.2	29.9	21.7	15.0	4.8	1.8
<=51	71.8	59.6	36.8	19.9	11.1	67.3	61.6	47.4	3.0	61.9	52.9	34.6	25.7	18.2	6.4	2.5
<=53	78.3	66.0	42.2	23.6	13.6	73.9	68.1	53.5	4.0	68.4	59.2	39.9	30.3	22.0	8.4	3.3
<=55	82.6	70.3	46.2	26.7	15.6	78.2	72.3	57.6	4.8	72.6	63.4	43.8	34.1	25.0	9.9	4.0
<=57	87.6	75.2	50.8	30.4	18.2	83.1	77.3	62.4	6.0	77.6	68.2	48.3	38.3	28.8	12.0	5.1
<=60	93.1	80.7	56.1	35.1	21.8	88.6	82.8	67.8	8.1	83.0	73.7	53.7	43.4	33.4	15.2	6.7
<=65	97.4	85.0	60.4	39.0	25.1	92.9	87.1	72.1	9.9	87.4	77.9	57.9	47.7	37.3	18.0	8.4
<=100	100.0	87.6	63.0	41.5	27.4	95.5	89.7	74.7	11.7	89.9	80.5	60.5	50.2	39.9	20.1	10.1

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 lines				Intl. 2011 PPP lines				Percentile-based lines						
		Food	100%	150%	200%	$1.90	$3.20	$5.50	$21.70	10th	20th	40th	50th	60th	80th	90th
<=25	4.8	86.0	62.6	41.4	27.4	92.6	88.1	74.0	11.7	88.4	79.6	60.2	50.2	39.9	20.1	10.1
<=29	9.7	82.8	61.2	40.5	27.2	88.4	84.5	71.9	11.7	84.8	76.9	58.8	49.3	39.5	19.9	10.1
<=32	14.8	79.9	60.1	40.1	27.0	84.0	81.2	70.0	11.7	81.5	74.7	57.9	48.7	39.1	19.9	10.1
<=35	20.9	74.9	57.8	38.9	26.5	78.1	75.9	66.7	11.6	76.1	70.7	55.7	47.5	38.3	19.7	10.0
<=37	25.1	71.7	56.4	38.1	26.0	74.0	72.5	64.7	11.6	72.5	68.2	54.5	46.6	37.6	19.5	10.0
<=39	31.0	66.8	53.8	37.1	25.6	68.4	67.4	61.2	11.5	67.4	64.1	52.0	45.1	36.8	19.2	9.9
<=41	37.3	61.1	50.8	35.6	24.8	62.2	61.6	56.8	11.4	61.6	59.2	49.3	43.2	35.5	18.8	9.8
<=43	44.2	54.8	46.5	33.3	23.6	55.5	55.1	51.2	11.1	55.1	53.3	45.3	40.1	33.6	18.4	9.7
<=45	51.6	47.7	41.8	30.9	22.2	48.1	47.9	45.3	10.8	47.9	46.7	40.8	36.9	30.9	17.7	9.4
<=47	59.2	40.4	36.4	27.8	20.3	40.7	40.5	38.8	10.3	40.5	39.7	35.5	32.8	28.2	16.7	9.0
<=49	65.8	33.8	31.1	24.8	18.4	34.0	33.9	32.8	9.5	33.9	33.3	30.6	28.6	24.8	15.3	8.3
<=51	71.8	28.0	26.2	21.6	16.3	28.2	28.1	27.2	8.7	28.1	27.6	25.9	24.6	21.7	13.7	7.6
<=53	78.3	21.5	20.8	17.9	13.8	21.6	21.6	21.2	7.7	21.6	21.3	20.7	19.9	17.9	11.7	6.8
<=55	82.6	17.3	16.8	14.8	11.8	17.3	17.3	17.1	6.9	17.3	17.2	16.7	16.1	14.9	10.2	6.1
<=57	87.6	12.4	12.2	11.1	9.1	12.4	12.4	12.3	5.7	12.4	12.3	12.2	11.9	11.1	8.1	5.0
<=60	93.1	6.9	6.8	6.4	5.5	6.9	6.9	6.8	3.6	6.9	6.8	6.8	6.8	6.5	4.9	3.4
<=65	97.4	2.6	2.6	2.5	2.3	2.6	2.6	2.6	1.8	2.6	2.6	2.6	2.6	2.5	2.1	1.7
<=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

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 lines				Intl. 2011 PPP lines				Percentile-based lines						
		Food	100%	150%	200%	$1.90	$3.20	$5.50	$21.70	10th	20th	40th	50th	60th	80th	90th
<=25	4.8	89.3	67.0	46.2	32.2	94.5	91.3	78.2	16.5	91.6	83.4	64.7	55.0	44.7	24.9	14.9
<=29	9.7	87.8	69.0	49.3	36.7	91.1	89.2	78.9	21.4	89.4	83.0	66.8	58.1	48.9	29.5	19.8
<=32	14.8	87.0	72.0	53.4	41.4	87.2	87.6	80.1	26.5	87.8	83.7	70.1	61.9	53.2	34.5	24.9
<=35	20.9	83.1	73.5	57.1	46.4	81.6	83.0	79.5	32.5	83.0	81.7	71.9	65.5	57.6	40.1	30.9
<=37	25.1	81.0	74.9	59.9	49.7	77.7	80.4	79.8	36.6	80.2	80.9	73.6	68.1	60.4	44.1	35.0
<=39	31.0	77.1	75.5	63.7	54.8	72.2	76.2	78.7	42.2	75.9	78.7	74.5	70.9	64.7	49.4	40.7
<=41	37.3	72.0	75.8	67.0	59.5	66.1	70.7	76.1	48.3	70.5	75.1	75.4	73.4	68.5	54.7	46.9
<=43	44.2	66.2	74.2	69.3	64.0	59.6	64.8	72.0	54.7	64.5	70.2	74.2	74.1	71.5	60.8	53.4
<=45	51.6	59.5	72.3	71.9	68.5	52.3	57.8	67.5	61.5	57.5	64.5	72.8	75.1	73.7	67.0	60.4
<=47	59.2	52.4	68.9	73.3	72.5	45.0	50.6	62.2	68.0	50.3	58.0	69.7	74.5	75.6	72.4	67.0
<=49	65.8	45.9	65.1	73.8	75.3	38.4	44.0	56.7	73.1	43.7	51.9	66.6	72.7	75.6	76.3	72.3
<=51	71.8	40.2	61.2	73.4	77.0	32.6	38.2	51.5	77.4	37.9	46.5	63.1	70.6	75.3	79.0	76.9
<=53	78.3	33.9	56.9	72.6	78.5	26.0	31.8	46.0	82.0	31.5	40.5	59.1	67.9	74.2	81.7	81.8
<=55	82.6	29.7	53.2	70.8	78.9	21.8	27.6	42.1	84.8	27.3	36.4	55.5	64.7	72.6	83.0	84.7
<=57	87.6	24.8	49.0	68.2	78.5	16.9	22.7	37.4	87.3	22.4	31.7	51.5	61.2	69.9	83.6	87.5
<=60	93.1	19.3	43.8	64.5	76.8	11.4	17.2	32.1	88.7	17.0	26.3	46.3	56.4	66.1	82.8	89.7
<=65	97.4	15.0	39.6	60.9	74.7	7.0	12.9	27.9	89.3	12.6	22.0	42.0	52.3	62.6	81.5	90.7
<=100	100.0	12.4	37.0	58.5	72.6	4.5	10.3	25.3	88.3	10.1	19.5	39.5	49.8	60.1	79.9	89.9

Scorecard applied to the validation sample.

For an example cut-off of 41 or less in the previous figures, 37.3 percent of all people are targeted, and outcomes for 100% of the national line in the validation sample are:

- Inclusion: 25.1 percent are below the line and correctly targeted
- Undercoverage: 11.9 percent are below the line and mistakenly not targeted
- Leakage: 12.2 percent are above the line and mistakenly targeted
- Exclusion: 50.8 percent are above the line and correctly not targeted

Increasing the cut-off to 43 or less increases the share of of all people targeted to 44.2 percent. The higher cut-off improves inclusion and undercoverage but worsens leakage and exclusion:

- Inclusion: 27.7 percent are below the line and correctly targeted
- Undercoverage: 9.3 percent are below the line and mistakenly not targeted
- Leakage: 16.5 percent are above the line and mistakenly targeted
- Exclusion: 46.5 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 41 or less is 75.8 percent. That is, about three in four Bolivians 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 lines				Intl. 2011 PPP lines				Percentile-based lines						
		Food	100%	150%	200%	$1.90	$3.20	$5.50	$21.70	10th	20th	40th	50th	60th	80th	90th
<=25	4.8	68.0	91.8	98.5	99.9	39.5	67.2	86.8	100.0	67.2	80.3	93.6	99.2	99.9	99.9	100.0
<=29	9.7	51.2	81.1	89.8	97.8	27.1	47.4	71.6	100.0	47.4	62.6	82.4	90.2	96.4	98.4	100.0
<=32	14.8	47.9	80.5	90.2	97.3	21.9	43.0	68.4	100.0	42.6	61.0	82.6	89.7	95.1	98.9	100.0
<=35	20.9	39.2	75.2	87.3	95.7	16.6	34.0	61.6	99.7	33.4	52.9	77.2	86.7	92.5	98.0	99.7
<=37	25.1	36.9	73.7	86.6	94.5	14.5	31.6	60.2	99.6	30.5	50.9	76.0	85.6	90.9	97.7	99.6
<=39	31.0	33.1	70.2	85.9	94.2	12.3	28.2	56.4	99.3	27.3	47.1	72.7	83.4	90.2	97.2	99.4
<=41	37.3	29.1	67.2	84.2	93.0	10.6	24.6	51.9	99.1	23.9	42.8	69.9	81.0	88.4	96.4	99.3
<=43	44.2	25.9	62.7	81.4	91.4	9.4	21.8	46.9	98.7	21.2	38.3	65.5	77.0	85.8	96.1	99.0
<=45	51.6	22.8	59.0	79.5	89.9	8.2	19.2	43.1	98.3	18.6	34.5	61.9	74.1	82.7	95.4	98.7
<=47	59.2	20.3	55.0	76.9	88.1	7.3	17.0	39.4	97.6	16.5	31.0	57.8	70.5	80.2	94.2	98.1
<=49	65.8	18.4	51.6	74.5	86.4	6.6	15.3	36.3	96.6	14.9	28.3	54.6	67.1	77.2	92.7	97.2
<=51	71.8	17.0	48.8	72.2	84.6	6.2	14.2	33.9	95.8	13.8	26.3	51.8	64.2	74.7	91.1	96.5
<=53	78.3	15.7	46.1	69.8	82.6	5.7	13.1	31.7	94.9	12.7	24.5	49.1	61.3	71.9	89.3	95.8
<=55	82.6	15.0	44.1	67.7	81.2	5.4	12.5	30.3	94.2	12.1	23.3	47.0	58.7	69.8	88.0	95.1
<=57	87.6	14.2	42.0	65.2	79.2	5.1	11.8	28.7	93.2	11.4	22.1	44.9	56.2	67.2	86.2	94.2
<=60	93.1	13.4	39.7	62.3	76.5	4.8	11.1	27.1	91.3	10.8	20.9	42.4	53.3	64.1	83.7	92.7
<=65	97.4	12.8	38.0	60.0	74.3	4.6	10.6	26.0	89.8	10.3	20.0	40.5	51.1	61.7	81.5	91.4
<=100	100.0	12.4	37.0	58.5	72.6	4.5	10.3	25.3	88.3	10.1	19.5	39.5	49.8	60.1	79.9	89.9

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 lines				Intl. 2011 PPP lines				Percentile-based lines						
		Food	100%	150%	200%	$1.90	$3.20	$5.50	$21.70	10th	20th	40th	50th	60th	80th	90th
<=25	4.8	2.1:1	11.2:1	66.9:1	813.3:1	0.7:1	2.0:1	6.6:1	All poor	2.0:1	4.1:1	14.7:1	131.3:1	813.3:1	813.3:1	All poor
<=29	9.7	1.0:1	4.3:1	8.8:1	43.7:1	0.4:1	0.9:1	2.5:1	All poor	0.9:1	1.7:1	4.7:1	9.2:1	26.6:1	60.2:1	All poor
<=32	14.8	0.9:1	4.1:1	9.2:1	36.4:1	0.3:1	0.8:1	2.2:1	All poor	0.7:1	1.6:1	4.7:1	8.7:1	19.2:1	86.1:1	All poor
<=35	20.9	0.6:1	3.0:1	6.9:1	22.1:1	0.2:1	0.5:1	1.6:1	363.3:1	0.5:1	1.1:1	3.4:1	6.5:1	12.3:1	47.8:1	363.3:1
<=37	25.1	0.6:1	2.8:1	6.5:1	17.0:1	0.2:1	0.5:1	1.5:1	223.0:1	0.4:1	1.0:1	3.2:1	5.9:1	10.0:1	43.0:1	223.0:1
<=39	31.0	0.5:1	2.4:1	6.1:1	16.4:1	0.1:1	0.4:1	1.3:1	135.0:1	0.4:1	0.9:1	2.7:1	5.0:1	9.2:1	35.0:1	157.8:1
<=41	37.3	0.4:1	2.1:1	5.3:1	13.4:1	0.1:1	0.3:1	1.1:1	115.5:1	0.3:1	0.7:1	2.3:1	4.3:1	7.6:1	26.5:1	135.5:1
<=43	44.2	0.3:1	1.7:1	4.4:1	10.7:1	0.1:1	0.3:1	0.9:1	76.4:1	0.3:1	0.6:1	1.9:1	3.4:1	6.1:1	24.6:1	97.0:1
<=45	51.6	0.3:1	1.4:1	3.9:1	8.9:1	0.1:1	0.2:1	0.8:1	56.9:1	0.2:1	0.5:1	1.6:1	2.9:1	4.8:1	20.6:1	76.9:1
<=47	59.2	0.3:1	1.2:1	3.3:1	7.4:1	0.1:1	0.2:1	0.7:1	40.3:1	0.2:1	0.4:1	1.4:1	2.4:1	4.1:1	16.1:1	50.7:1
<=49	65.8	0.2:1	1.1:1	2.9:1	6.3:1	0.1:1	0.2:1	0.6:1	28.7:1	0.2:1	0.4:1	1.2:1	2.0:1	3.4:1	12.7:1	35.1:1
<=51	71.8	0.2:1	1.0:1	2.6:1	5.5:1	0.1:1	0.2:1	0.5:1	22.5:1	0.2:1	0.4:1	1.1:1	1.8:1	2.9:1	10.2:1	27.6:1
<=53	78.3	0.2:1	0.9:1	2.3:1	4.8:1	0.1:1	0.2:1	0.5:1	18.5:1	0.1:1	0.3:1	1.0:1	1.6:1	2.6:1	8.3:1	22.5:1
<=55	82.6	0.2:1	0.8:1	2.1:1	4.3:1	0.1:1	0.1:1	0.4:1	16.2:1	0.1:1	0.3:1	0.9:1	1.4:1	2.3:1	7.4:1	19.5:1
<=57	87.6	0.2:1	0.7:1	1.9:1	3.8:1	0.1:1	0.1:1	0.4:1	13.6:1	0.1:1	0.3:1	0.8:1	1.3:1	2.0:1	6.3:1	16.2:1
<=60	93.1	0.2:1	0.7:1	1.7:1	3.3:1	0.1:1	0.1:1	0.4:1	10.5:1	0.1:1	0.3:1	0.7:1	1.1:1	1.8:1	5.1:1	12.8:1
<=65	97.4	0.1:1	0.6:1	1.5:1	2.9:1	0.0:1	0.1:1	0.4:1	8.8:1	0.1:1	0.3:1	0.7:1	1.0:1	1.6:1	4.4:1	10.6:1
<=100	100.0	0.1:1	0.6:1	1.4:1	2.7:1	0.0:1	0.1:1	0.3:1	7.5:1	0.1:1	0.2:1	0.7:1	1.0:1	1.5:1	4.0:1	8.9:1

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 lines				Intl. 2011 PPP lines				Percentile-based lines						
		Food	100%	150%	200%	$1.90	$3.20	$5.50	$21.70	10th	20th	40th	50th	60th	80th	90th
<=25	4.8	26.4	12.0	8.1	6.6	42.4	31.4	16.6	5.5	32.3	19.9	11.4	9.6	8.0	6.0	5.4
<=29	9.7	40.0	21.3	14.9	13.1	58.7	44.6	27.5	11.0	45.9	31.3	20.3	17.6	15.6	12.0	10.8
<=32	14.8	56.9	32.1	22.8	19.8	72.0	61.4	39.9	16.7	62.6	46.2	30.9	26.6	23.3	18.3	16.4
<=35	20.9	65.8	42.4	31.1	27.5	77.1	68.7	50.8	23.6	69.4	56.7	40.8	36.4	32.1	25.6	23.2
<=37	25.1	74.5	50.1	37.2	32.7	81.0	76.9	59.8	28.3	76.4	65.6	48.3	43.2	38.0	30.8	27.8
<=39	31.0	82.4	58.7	45.5	40.2	84.9	84.4	69.0	34.8	84.1	74.8	57.0	51.9	46.4	37.7	34.2
<=41	37.3	87.1	67.7	53.7	47.8	88.1	88.8	76.5	41.9	88.7	81.9	66.0	60.7	54.8	45.0	41.2
<=43	44.2	91.9	74.8	61.5	55.6	91.9	93.2	81.9	49.4	93.0	86.8	73.3	68.4	63.0	53.1	48.7
<=45	51.6	94.6	82.3	70.1	63.9	93.8	95.6	87.8	57.4	95.5	91.4	80.8	76.9	71.0	61.6	56.7
<=47	59.2	96.6	87.9	77.8	71.8	96.6	97.1	92.1	65.4	97.0	94.1	86.6	83.8	78.9	69.8	64.6
<=49	65.8	97.2	91.8	83.9	78.3	96.9	97.4	94.5	72.0	97.4	95.6	91.0	88.7	84.5	76.4	71.2
<=51	71.8	98.0	94.6	88.6	83.6	98.3	98.2	96.0	77.8	98.2	96.9	94.1	92.6	89.1	81.8	77.1
<=53	78.3	99.1	97.7	93.5	89.2	98.9	99.0	98.1	84.2	99.0	98.3	97.5	96.5	93.7	87.6	83.4
<=55	82.6	99.6	98.4	95.7	92.4	99.4	99.6	98.9	88.2	99.6	98.9	98.3	97.5	95.9	91.0	87.4
<=57	87.6	99.7	99.4	97.7	95.5	99.4	99.7	99.4	92.4	99.7	99.4	99.5	99.0	97.8	94.5	91.8
<=60	93.1	99.9	99.8	99.2	98.1	99.8	99.9	99.7	96.3	99.9	99.6	99.8	99.8	99.2	97.5	96.0
<=65	97.4	99.9	100.0	99.9	99.6	99.8	99.9	100.0	99.1	99.9	100.0	100.0	100.0	99.9	99.4	99.0
<=100	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

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 39 or less would target 31.0 percent of Bolivians and give a head-count poverty rate among those targeted of 70.2 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 39 or less, about 2.4 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 Bolivians. For 100% of the national line, **Figure 21** shows that a score cut-off of 37 or less would target 25.1 percent of all Bolivians, a segment that includes 50.1 percent of all poor Bolivians.

Interview Guide

Citations in the **Interview Guide** come from:

Instituto Nacional de la Estadística. (2019) “Manual del/la Encuestador(a) : EH 2019” [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 exact 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 department does the household live?”). Instead, fill in the response based on the knowledge that you the enumerator have of the department 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 (“How many household members work in their main occupation as wage/salary employees, domestic servants, or business owners who draw a salary?”). 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**.

Finally, do not directly ask the fourth scorecard question ("How many household members in their main occupation are self-employed, business owners who do not draw a salary, or members of a producer's cooperative?"). 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 all of the six 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 such 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 is unknown, does not exist, or is not applicable, then write "UNKNOWN" or "None" 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 collect data that cannot be obtained.

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. Is the female head (or spouse of the male head) covered by medical insurance?	A. No	0	
	B. Yes	2	**2**
	C. No female head (nor spouse of the male head)	4	

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 Bolivia's INE in the 2019 EH. 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 2019 EH by Bolivia's INE. For example, interviews should done in-person by a trained enumerator at the residence of the participating household because that is what INE did in the 2019 EH.

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 **Interview Guide** are available in **Spanish** and **English**.There are not yet official, professional translations to other languages spoken in Bolivia such as Quechua and Aymara. 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*

G3.1 Do's

According to pp. 35−37 of the *Manual*, you the enumerator should:

- "Study [this **Interview Guide**] carefully so as to understand it completely and to perform your work efficiently
- Always carry your badge that identifies you as employed by your program, your personal identification, [this **Interview Guide**], and all other materials required to do your work
- Check to make sure that you have asked all the questions [in the scorecard questionnaire] and that you have recorded responses for all of them
- Avoid—to the extent possible—interviewing in the presence of third parties who are not members of the interviewed household

G3.2 Don'ts

- Show up for work dressed inappropropriately or in a mental or physical state that is not conducive to excellent work
- Bring along to an interview a third party who has no business being there
- Delegate your tasks to unauthorized people
- Make up or assume responses to questions
- Divulge any household's responses, as that would violate confidentiality
- Argue with respondents
- Ask questions unrelated to the interview
- Promise rewards or incentives in return for the respondent's cooperation

G3.3 Greeting the household to be interviewed

According to pp. 38−42 of the *Manual*, an interview has three parts : introduction, the interview proper, and the wrap up.

The *introduction* "is your first meeting with the members of the household to be interviewed. When done well, it goes a long way towards the eventual completion of a successful interview. You the enumerator should:

- "Dress appropriately for the area where you are working. This will encourage respect, trust, and ultimately cooperation from the respondent
- Introduce yourself by stating your name and showing your badge to establish that you are an agent of [your program]
- Greet the household warmly, and do the interview respectfully, always building an atmophere of trust
- If the household does not wish to cooperate, then politely persist. Explain that the information collected will be kept strictly confidential and will be used only for statistical purposes
- Explain the reason for your visit simply and clearly, making it easy for the respondent to understand. Carefully read the following example which may serve a template for your own introduction:

  > "Good morning. My name is [YOUR NAME]. I work with [YOUR PROGRAM]. [Here is my badge.] Your [household] was selected for a short interview that is meant to help our program get to know [our participants] better. For example, we would like to know how many [members are in your household], their ages, what work they do, and so on. The goal is [to improve the performance of our program].

"After introducing yourself and building a climate of trust, you the enumerator should be sure to maintain this trust throughout the interview. The following strategies will help you to manage the interview successfully:

Attention: "Pay attention at all times. This shows respect and courtesy to the respondent, encouraging him/her to reciprocate. Paying close, constant attention will also help you the enumerator to collect high-quality data and to avoid losing track of your place in the sequence of questions.

Pace. “Do not treat the respondent like a ‘response machine’, as this can lead to inaccurate responses. Try to read the questions at a steady speed; do not start slow and speed up, and do not start fast and then slow down. Figure out how quickly the respondent catches on, and then adjust how quickly you ask the questions accordingly. Read each word of each question clearly. If you notice that the respondent is getting tired or annoyed, remind him/her that the interview will not take long, and maintain a good rhythm. Try not to take long breaks. If you need to pause to chat off-topic with a tired or distracted respondent, then be brief and try to steer the conversation back around to the interview.

Self-control. “Sometimes, you the enumerator will have to deal with unexpected inconveniences (for example, interrupting children, visiting salespeople, and ringing telephones). Remind yourself that these nuisances are part of the every-day life of the household and that the household itself is used to them and thus does not perceive them as annoying. Therefore, be patient and prudent, keeping your emotions in check. As you wait to continue the interview, do not show any impatience or anger. It is better that the respondent feels trusted and responds willingly and gratefully than to refuse to participate because you seem intolerant.

Focus on the work. In the course of an interview, you may be left alone for a time, for example, because the respondent gets up to attend to something in the kitchen or because he/she takes a phone call. At such times, do not get up and wander around nor look at objects or peek at papers that may be lying about the room. This is impolite and in bad taste. After all, not only are you a visitor, but you also invited yourself in. Stay calm, stay seated, and focus your thoughts on what you can do to help the rest of the interview go well. Take advantage of the chance to review the responses so far and to plan the rest of the work so that you do not have to come back to the household again later.

Interview management. “Sometimes, a respondent will give obviously inaccurate or irrelevant responses, make side—or snide—comments, or simply beat around the bush. When this happens, do not suddenly or rudely cut him/her off. Instead, listen calmly to what he/she has to say, and then gently guide the discussion back to the questions in the scorecard questionnaire.

Neutrality. “The scorecard questions are carefully crafted to be neutral and to avoid suggesting that any particular response is better or preferred. Likewise, you the enumerator must maintain complete and constant neutrality during the interview. Do not allow yourself to show any surprise, approval, or disapproval about anything that the respondent says, whether by your facial expression, tone of voice, or body language. Nor should you ever express a personal opinion. If the respondent asks for your opinion, explain that, to avoid influencing the results of the survey, you would be happy to discuss your personal views after the interview is complete.

Probing. “If a respondent gives a vague or ambiguous response, do not suppose that he/she did this on purpose, and do not be taken aback. Simply ask probing questions for clarification.

Clarify questions. “If the respondent says that he/she does not understand a question, then explain what the question means.

Clarify responses: “If the respondent’s statements are confusing, or if you the enumerator do not understand them, then ask for clarification or for complementary information.

“In the wrap-up of the interview, do your best to leave a positive impression of yourself, the work you have done, and [your program].

“Once the interview is complete, review the responses that you have recorded, making sure that everything is complete and correct. If you do this before you take your leave of the interviewed household, then you can fill in any gaps while your are still with the respondent.

Finally, take your leave graciously, thanking the respondent and the other members of the interviewed household profusely for their cooperation. Be sure to let them know that you may return in the future to ask them for additional help if it turns out to be needed.

G3.4 Who should be the respondent?

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

According to p. 26 of the *Manual*, "The preferred respondent is the head of the household (whether male or female). If he/she is not available, then the respondent can be any household member who is at least 12-years-old."

G3.5 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 because no one can be a member of more than one household.

According to p. 21 of the *Manual*, the *head of the household* is "the person recognized as the such by the rest of the household members, regardless of the head's age, sex, or marital status. If the members of the household cannot come to an agreement or are otherwise unable to identify their head, then ask the respondent to decide based on his/her answers to two questions: 'Who is responsible for the well-being of the household?' and 'Who makes the key decisions?'."

G4. Guidelines for each question in the scorecard

G4.1 In which department does the household live?

A. Chuquisaca
B. La Paz
C. Tarija, or Oruro
D. Cochabamba, or Potosí
E. Pando, or Beni
F. Santa Cruz

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

G4.2 How many members does the household have?

A. Six or more
B. Five
C. Four
D. Three
E. Two
F. One

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

According to p. 19 of the *Manual*, a *household* is "a unit of people (regardless of blood or marital relationship) who live together in the same residence and who meet their basic needs from a shared budget. That is, members of the household share expenses, regardless of who provides the resources to meet those expenses. A single person living alone is a one-person household."

According to pp. 78−79 of the *Manual*, "The following count as *members of the interviewed household*:

- "People who normally live together in a residence and who share a budget (regardless of whether any given member contributes to the budget).
- People who normally live in the residence and who share a budget but who, at the time of the interview, are temporarily absent for a period whose total actual or expected duration is three months or less. The absence may be due, for example, to vacation, visits with friends or relatives, business trips, hospitalization, and so on

"The following do not count as *members of the interviewed household*:

- People staying temporarily in the residence but who usually live elsewhere. This includes, for example:
  - Students (or members of the armed forces) who usually live elsewhere (boarding schools, residential colleges, military schools, or barracks)
  - Members of the armed forces who normally live in military housing
  - People who contribute financially to the household but who—for reasons of work or study—usually live elsewhere
- Foreigners on official business in Bolivia (for example, embassadors or consuls), as well as other foreigners who will stay in the residence only for a short time
- Children of household members who have been absent for more than three months—whether for work, school, or other reasons—regardless of whether those children depend economically on the interviewed household
- Lodgers who usually live in the residence with the interviewed household but who pay for their room and board"

According to p. 80 of the *Manual*, you the enumerator should count as members of the interviewed household "anyone who usually lives in the residence and who has been absent for less than three months or who—despite not having lived with the household for at least three months—plans to be a permanent part of the household from now on.

"Count people who are temporarily absent such as children, newborns, the elderly, visitors (as long as they usually live in the residence and plan to continue), and domestic servants who do not have another residence elsewhere (and the families of domestic servants) [as long as these people also meet all the other criteria to be members of the interviewed household].

"Record the members of the household in this order:

- Head
- Spouse/conjugal partner of the head
- Unmarried children of the head (starting with the oldest)
- Married children of the head, their spouses, and their children
- Parents and in-laws of the head, other relatives, and other non-relatives, concluding with any live-in domestic servants and their relatives"

G4.3 How many household members work in their main occupation as wage/salary employees, domestic servants, or business owners who draw a salary?

A. None
B. One
C. Two or more

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 in the second-to-last column of the **Back-page Worksheet**.

This question concerns the work status only of members of the interviewed household who (based on the data from the **Back-page Worksheet**):

- Are seven-years-old or older, and
- Worked at least one hour in the past calendar-week

According to p. 137 of the *Manual*, *work* is "any economic activity done in exchange for pay (whether in-cash or in-kind), including activities that help earn income or that contribute to an economic activity run by the household, even if those activities are not directly productive or directly remunerated.

"A person is considered to have worked if he/she did such an activity for at least one hour in the past calendar-week."

"Keep in mind the following criteria for determining whether an activity is to be considered as *work*.

"Activities that count as *work*:

- Activities done for an organization, institution, business, or office (public or private)
- Sale of items in the street, whether from a fixed location or itinerantly
- Provision of domestic services to other households. Examples include work as a domestic servant, child care, cooking, cleaning, and so on

"Work in a household business or as an unpaid apprentice/intern counts as *work*, even though it is not remunerated directly.

“Activities that do not count as *work*:

- Domestic chores done by members of the household who are unpaid and unremunerated
- Volunteer work in the community that is unpaid and unremunerated
- Unpaid and unremunerated work done by a member of the household for a family member who has a salaried job (for example, helping to type a text for a brother who is in the civil service)
- Activities such as personal hygiene, recreation, and other self-maintenance activities such as eating and sleeping
- Begging or theft”

According to page 146 of the *Manual*, a person’s *main occupation* is “the one in which the person earned the most income in the reference period. If a person has two or more occupations, and if the person earned the same income from all of them in the reference period, then the main occupation is the one in which the person worked for the most time. If the person worked the same time in all of the occupations, then you the enumerator should ask the respondent to choose one as the main occupation.”

According to pp. 159 and 160 of the *Manual*, a *wage/salary employee* is “a person who performs physical work or skilled/knowlege work for an employer (be it public or private) in return for a wage or salary (be it in-cash or in-kind).

A *domestic servant* is “a person who works for a household in its residence doing domestic chores in return for a wage or salary (in-cash or in-kind). This includes household servants of any type who have an employment contract (regardless of whether they stay overnight in the residence of the household that employs them). Examples include nannies, nurses, cooks, butlers, gardeners, and so on.

A *business owner who draws a salary* is “a person who runs his/her own business or economic activity and who employs wage/salary employees. This type of business owner draws a monthly salary for the work he/she does in his/her own business. In addition, the business owner has the rights to some share of the profits of the business. Businesses with this type of owner usually keep formal accounting records and have formal financial controls in place.”

According to p. 136 of the *Manual*, "The reference period is the calendar-week just prior to the day of the interview.

"For example, if the interview is on Saturday, March 26, 2022, then the reference period is from Monday, March 14 to Sunday, March 20 (inclusive). [The reference period does not run from Saturday, March 19 to Friday, March 25, nor does the reference period run from Monday, March 21 to Friday, March 25.] In other words, the reference period is not the current calender-week, but rather the calendar week starting on Monday of the previous calendar-week."

March 2022						
Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
Calender-week for the reference period						
21	22	23	24	25	26	27
28	29	30	31	Day of the interview		

G4.4 How many household members in their main occupation are self-employed, business owners who do not draw a salary, or members of a producer's cooperative?

A. None
B. One
C. Two or more

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 in the last column of the **Back-page Worksheet**.

This question concerns the work status only of members of the interviewed household who (based on the data from the **Back-page Worksheet**):

- Are seven-years-old or older, and
- Worked at least one hour in the past calendar-week

According to p. 137 of the *Manual*, *work* is "any economic activity done in exchange for pay (whether in-cash or in-kind), including activities that help earn income or that contribute to an economic activity run by the household, even if those activities are not directly productive or directly remunerated.

"A person is considered to have worked if he/she did such an activity for at least one hour in the past calendar-week."

"Keep in mind the following criteria for determining whether an activity is to be considered as *work*.

"Activities that count as *work*:

- Activities done for an organization, institution, business, or office (public or private)
- Sale of items in the street, whether from a fixed location or itinerantly
- Provision of domestic services to other households. Examples include work as a domestic servant, child care, cooking, cleaning, and so on

"Work in a household business or as an unpaid apprentice/intern counts as *work*, even though it is not remunerated directly.

“Activities that do not count as *work*:

- Domestic chores done by members of the household who are unpaid and unremunerated
- Volunteer work in the community that is unpaid and unremunerated
- Unpaid and unremunerated work done by a member of the household for a family member who has a salaried job (for example, helping to type a text for a brother who is in the civil service)
- Activities such as personal hygiene, recreation, and other self-maintenance activities such as eating and sleeping
- Begging or theft”

According to page 146 of the *Manual*, a person’s *main occupation* is “the one in which the person earned the most income in the reference period. If a person has two or more occupations, and if the person earned the same income from all of them in the reference period, then the main occupation is the one in which the person worked for the most time. If the person worked the same time in all of the occupations, then you the enumerator should ask the respondent to choose one as the main occupation.”

According to pp. 159 and 160 of the *Manual*, a *self-employed worker* is “a person who has his/her own business or economic activity. A self-employed worker does not work for someone else, and he/she does not employ any wage/salary employees. A self-employed worker buys and sells, produces goods for sale, or provides services. He/she may or may not have help from unremunerated household workers or interns/apprentices. Examples include tailors, carpenters, travelling salespeople, plumbers, and so on.

”The category of self-employed workers also includes people who work for others by the task, without a permanent contract for their general labor or for multiple tasks whose nature is not pre-specified. In this arrangement, the self-employed worker receives a pre-determined remuneration in return for performing a pre-specified task. The relationship ends when the task is completed. Examples include bricklayers and concrete workers who work by the task, painters, electricians, tailors, gardeners, clothes-washers, and so on.

“The category of self-employed workers also includes people who work in their own residence or in the residences of other people (according to the worker’s choice) by the task or by the order. They are not directly managed or controlled by their clients, so they choose how to do their own work; the clients care only for the finished task or order. Other than the payment contracted for the task or order,

these self-employed workers do not receive benefits from their clients nor are they covered by unemployment insurance from the government.

A *business owner who does not draw a salary* is "a person who runs his/her own business or economic activity and who employs wage/salary employees. Businesses with this type of owner usually do not keep formal accounting records and do not have formal financial controls in place.

A *member of a producer's cooperative* is "a share-holding member of a cooperative who actively works in the cooperative's business and who receives income and absorbs losses as a member of the cooperative. Examples include cooperative miners, gold miners, and so on."

According to p. 136 of the *Manual*, "The reference period is the calendar-week just prior to the day of the interview.

"For example, if the interview is on Saturday, March 26, 2022, then the reference period is from Monday, March 14 to Sunday, March 20 (inclusive). [The reference period does not run from Saturday, March 19 to Friday, March 25, nor does the reference period run from Monday, March 21 to Friday, March 25.] In other words, the reference period is not the current calender-week, but rather the calendar week starting on Monday of the previous calendar-week."

March 2022						
Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
Calender-week for the reference period						
21	22	23	24	25	26	27
28	29	30	31	Day of the interview		

G4.5 Is the female head (or spouse of the male head) covered by medical insurance?

A. No
B. Yes
C. No female head (nor spouse of the male head)

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 because no one can be a member of more than one household.

According to p. 21 of the *Manual*, the *head of the household* is "the person recognized as the such by the rest of the household members, regardless of the head's age, sex, or marital status. If the members of the household cannot come to an agreement or are otherwise unable to identify the head, then ask the respondent to decide based on his/her answers to two questions: 'Who is responsible for the well-being of the household?' and 'Who makes the key decisions?'."

Remember that you the enumerator already know the name of the female head (or spouse of the male head) and whether she exists from compiling the **Back-page Worksheet**. Thus, if there is a female head (or a spouse of the male head), then do not mechanically ask, "Is the female head (or spouse of the male head) covered by medical insurance?" Instead, use the actual name of the female head (or spouse of the female head), for example: "Is María covered by medical insurance?"

If there is no female head (nor spouse of the male head), then do not ask the question of the respondent at all. Instead, mark "C. No female head (nor spouse of the male head)" and go to the next question.

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

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

G4.6 How many rooms does the household occupy, not counting bathrooms, kitchen, laundry rooms, garages, or rooms used for storage or business?

A. One
B. Two or three
C. Four or more

According to page 75 of the *Manual*, a *room* is "any physical space with a roof and walls that can accommodate an adult-sized bed or cot and that is used as shelter by people. Example include bedrooms, dining rooms, and so on.

"If a physical space is demarcated by a piece of furniture or a curtain so as to establish its division into different uses, then it is to be counted as a single room (not as two rooms), even if the respondent considers the space to be two rooms. Furthermore, do not count rooms that serve economic functions, such as for running a business or for the provision of services, and so on."

According to p. 19 of the *Manual*, a *residence* is "a building with one or more floors covered by a roof that was designed or adapted to be temporarily or permanently used as shelter for one or more people. A residence has direct and unimpeded access from the street or has access through public spaces such as corridors, courtyards, or stairs."

G4.7 Is firewood, dung, manure, or llama pellets the main fuel or source of energy used for cooking?

A. Yes
B. No

According to page 74 of the *Manual*, the concept of *firewood* includes "split wood, logs, branches of bushes, and *yareta*.

"*Dung*, *manure*, or *llama pellets* refers to the dried excrement of llamas, sheep, goats, cattle, and so on (called *taquía* en the Altiplano).

G4.8 Does the household possess and use a frigerator or freezer?

A. No
B. Yes

According to p. 219 of the *Manual*, record "B. Yes" only if the interviewed household:

- Has a refrigerador or freezer in its possession, and
- Currently uses the refrigerador or freezer

If either requirement is not met, then record "A. No".

If the interviewed household reports possessing a refrigerador or freezer, then you the enumerator should probe further to determine whether the item is currently in use. For example, perhaps the interviewed household possesses a refrigerator but does not currently use it because it is out-of-order.

Possession does not necessarily imply *ownership*. The interviewed household might own a refrigerator without possessing it because it was lent out or rented out to another household for that other household's use. In this case, the response to be recorded is "A. No".

It is also possible that the interviewed household possesses a refrigerator and currently uses it even though it does not own the refrigerator but rather has borrowed it in, rented it in, or bought it with a loan that is not yet paid-off. In this case, both criteria for recording "B. Yes" have been met.

G4.9 Does the household possess and use an old-style tube TV or one or more flat-screen TVs (plasma, LCD, LED)?

A. No
B. Only tube
C. One flat screen (regardless of tube)
D. Two or more flat screens (regardless of tube)

According to p. 219 of the *Manual*, count an old-style tube TV or a flat-screen TV only if the interviewed household:

- Has the TV in its possession, and
- Currently uses the TV

If either requirement is not met, then do not count the TV.

If the interviewed household reports possessing a TV, then you the enumerator should probe further to determine whether the TV is currently in use. For example, perhaps the interviewed household possesses a TV but does not currently use it because it is out-of-order. Do not count TVs that the interviewed household does not currently use.

Possession does not necessarily imply *ownership*. The interviewed household might own a TV without possessing it because it was lent out or rented out to another household for that other household's use. Do not count such a TV even though the interviewed household owns it.

It is also possible that the interviewed household possesses a TV and currently uses it even though it does not own the TV but rather has borrowed it in, rented it in, or bought it with a loan that is not yet paid-off. In this case, both criteria have been met and so the TV should be counted even though the interviewed household does not own it.

You the enumerator should break this question into two questions, each asked of the respondent:

- Does the household possess and use an old-style tube TV?
- How many flat-screen TVs (plasma, LCD, LED) does the household possess and use?

Mark the response on the scorecard questionnaire according to the combination the two responses to these two questions as follows:

Any old-style tube TVs?	**How many flat-screen TVs?**	**Response to mark**
No	None	A
No	One	C
No	Two or more	D
Yes	None	B
Yes	One	C
Yes	Two or more	D

G4.10 How many household members have a cell phone available for their own personal (non-business) use?

A. None
B. One
C. Two
D. Three or more

According to p. 132 of the *Manual*, "A person can have a cell phone for his/her personal (non-business) use even if he/she does not own the phone or does not pay for the costs of its services. The key criterion is that the person has a phone available for his/her personal (non-business) use—within reasonable limits—thanks to his/her job, a friend, a relative, or some other arrangement. A person is not considered to have a cell phone available for his/her personal (non-business) use if the phone is used only occasionally, for example, by borrowing a cell phone from someone else temporarily in order to make a call."

Technical Annexes: Overview

The technical annexes cover advanced or technical 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

Bolivia's Instituto Nacional de Estadística (INE) fielded the 2019 Encuesta de Hogares (EH, the Household Survey) with 11,869 households from October 19 to December 20, 2019. The 2019 EH is Bolivia's most-recent national household income survey that collects data on possession of consumer durables.

Questions and points for the scorecard are selected (*constructed*) based on data from a random three-fifths of the 11,639 households in the 2019 EH for whom INE reports income.[34] These same three-fifths of households are also used to associate (*calibrate*) scores with poverty likelihoods for all supported poverty lines.

Data from the other two-fifths of households from the 2019 EH 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.

[34] INE does not report income for the 30 households with live-in domestic servants.

Annex 2 Definition of *poverty*

A household's *poverty status* as poor or non-poor depends on whether its income (BOB per person per day) is below a given poverty line. Thus, a definition of *poverty* is a poverty line together with a measure of income from the 2019 EH.

INE (pp. 12 and 67, **2020a**) defines income for the 2019 EH.

Because pro-poor programs in Bolivia may want to use different or various poverty lines, the scorecard supports 15 lines:

- Food (extreme) line
- 100% of the national line
- 150% of the national line
- 200% of the 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
- First-decile (10th-percentile) line
- First-quintile (20th-percentile) line
- Second-quintile (40th-percentile) line
- Median (50th-percentile) line
- Third-quintile (60th-percentile) line
- Fourth-quintile (80th-percentile) line
- Tenth-decile (90th-percentile) line

A2.1 National poverty lines

Bolivia uses the cost-of-basic-needs approach[35] to derive its two official poverty lines: a food (extreme) line, and a food-plus-non-food line.

[35] **Ravallion**, 1998.

A2.1.1 Food (extreme) line

Bolivia's food poverty line (*línea extrema*) is the cost of of a food basket that provides a minimum standard of 2,097 Calories per person per day in urban areas and 2,276 Calories in rural areas.

There are three food baskets based on the patterns of food consumption expenditure observed in Bolivia's 2015/16 Household Budget Survey (*Encuesta de Presupuestos de Familias*, EPF) for each of three types of areas: El Alto and the nine departmental capitals; other urban areas; and rural areas.

The food (extreme) line is the cost of a food basket, calculated separately for each of the 12 units in the three types of food-basket areas just described. Food-price indexes are used to update the 12 area-specific food lines from prices during the 2015/16 EPF to prices during the 2019 EH.

The all-Bolivia average food line in prices for Bolivia overall during the 2019 EH is about BOB14 per person per day, giving an all-Bolivia head-count poverty rate for this line of 12.9 percent (**Figure 10**).

A2.1.2 National line

For each of the 12 poverty-line regions, the national poverty line (usually called here "100% of the national line") is the region's food line, plus a minimum standard for non-food consumption expenditure. This non-food standard is defined separately for the three food-basket areas as the average observed non-food consumption expenditure for households in the 2015/16 EPF whose food consumption expenditure is close to the food line. After adjusting prices for the 12 regional non-food standards from the 2015/16 EPF to the 2019 EH, the average national (food-plus-non-food) poverty line for Bolivia overall is about BOB28 per person per day, with a head-count poverty rate of 37.2 percent (**Figure 10**).[36]

150% of the national line and 200% of the national line are multiples of 100% of the national line.

[36] The all-Bolivia head-count poverty rates reported here for the food line and 100% of the national line match those in **INE** (2020a, p. 74), suggesting that this paper uses the same data and calculations as INE did.

A2.2 International 2011 PPP poverty lines

The World Bank tracks world-wide poverty with four 2011 PPP poverty lines:[37]

- $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

The purpose of PPP lines is to 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 international comparability of poverty estimates.

International 2011 PPP lines for Bolivia are derived from:

- 2011 PPP (revised) exchange rate for Bolivia for "individual consumption expenditure by households":[38] BOB2.891 per $1.00
- Average all-Bolivia Consumer Price Index[39] (CPI):
  - Calendar-year 2011: 109.884
  - October 19 to December 20, 2019: 150.157
- Average person-weighted geographic price index across Bolivia's 12 poverty-line regions: 1.0000
- Geographic price indexes for each of Bolivia's 12 poverty-line regions:
  - Rural: 0.79700
  - Other urban: 0.94211
  - Santa Cruz (capital of Santa Cruz): 0.98130
  - Trinidad (capital of Beni): 1.02910
  - Oruro (capital of Oruro): 1.05406
  - Cobija (capital of Pando): 1.11069
  - Potosí (capital of Potosí): 1.12065
  - Sucre (capital of Chiquisaca): 1.21717
  - Cochabamba (capital of Cochabamba): 1.24335
  - El Alto: 1.24993
  - Tarija (capital of Tarija): 1.25108
  - La Paz (capital of La Paz): 1.26900

[37] **Jolliffe and Prydz**, 2016; **Ferreira *et al.***, 2016.
[38] **World Bank**, 2020, Table E.3, column 13, p. 134.
[39] Base = 100 in calendar-year 2010, **link**.

Given this, the $5.50/day 2011 PPP line for Boliva overall is:

$$\$5.50 \cdot 2011 \text{ PPP factor} \cdot \frac{CPI_{2019\,EH}}{CPI_{2011}} \cdot \frac{CPI_{region}}{CPI_{ave.}} = \$5.50 \cdot 2.891 \cdot \frac{150.157}{109.884} \cdot \frac{1.000}{1.000} = \text{BOB}21.73.$$

The corresponding head-count poverty rate is 25.7 percent (**Figure 10**).

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

The **World Bank** classifies Bolivia as an lower-middle income country. The corresponding $3.20/day line of about BOB13 per person per day is lower than Bolivia's food (extreme) line. Because the food line is stingy (people below this line are going hungry), $5.50/day is more appropriate for Bolivia.

The $5.50/day line for Bolivia of BOB21.73 and its head-count poverty rate of 25.7 percent differ from those of the World Bank's **PovcalNet** (BOB21.89 and 19.9 percent). Both use the 2019 EH and the same CPI series, so the difference are likely due to:

- Different average CPIs for calendar-year 2011 or for the 2019 EH, or
- PovcalNet's possible use of a single line for all-Bolivia, without adjusting for differences in prices in 12 regions (as done here and by Bolivia's INE)

The average all-Bolivia lines differ by 0.7 percent, a small difference vis-à-vis the 5.8-percentage-point difference in poverty rates. Thus, the most-likely explanation is that PovcalNet did not make regional price adjustments.[40]

Which 2011 PPP lines are to be preferred? If it makes sense to adjust for price differences across countries (the purpose of international PPP lines in the first place), then it also makes sense to adjust for price differences within a country. In addition, the documentation of the lines here is more complete (and thus more amenable to critique and improvement) than that of PovcalNet. Both factors favor this paper's 2011 international PPP lines and poverty rates.

[40] PovcalNet usually adjusts for regional prices when the data provided by a country's statistical authority has the adjustment baked into income. If the adjustments are made instead in the poverty lines (as in Bolivia), then PovcalNet is more likely to ignore them.

A2.3 Percentile-based poverty lines

The scorecard supports percentile-based poverty lines.[41] This facilitates a number of types of analyses. For example, the second-quintile (40th-percentile) line might be used to help track Bolivia'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 supported percentile lines) can also be used to look at the relationship of income with health outcomes (or anything else related with the distribution of income). 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.[42]

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

- Relative wealth (via scores)
- Absolute income (via poverty likelihoods and absolute poverty lines)
- Relative income (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, income 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*.

---

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

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

Annex 3 Scorecard construction

For Bolivia, about 65 candidate questions are prepared in these areas:

- Household composition (such as the number of household members)
- Education (such as whether the female head—or the spouse of the male head—can read and write))
- Employment (such as the number of household members who work as wage/salary employees, domestic servants, or business owners who draw a salary)
- Health (such as whether the female head—or the spouse of the male head—is covered by medical insurance)
- Housing (such as the number of rooms or the main fuel used for cooking)
- Ownership of consumer durables (such as TVs or cell phones)
- Food security (such as whether the household ever ran out of food in the past 12 months)
- Agriculture (such as the number of household members who work in agriculture)
- Location of residence (such as the department)

To facilitate the estimation of change over time, preference is given to questions with greater sensitivity 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 100% 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.[43]

[43] **Ravallion**, 2009.

One of the one-question draft scorecards is then selected based on:[44]

- Improvement in accuracy
- Acceptability to users in terms of:
  - Simplicity
  - Cost of collection
  - Concordance with:
    - Experience
    - Theory
    - Common sense
- Sensitivity to changes in income
- Variety among types of questions
- Applicability across departments
- Tendency to have a slow-changing relationship with poverty
- Relevance for distinguishing among people at the poorer end of the distribution of income
- 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 step. 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[45] and non-statistical criteria. The use of non-statistical criteria can improve robustness against violations in the scorecard's basic assumptions. It also helps to ensure that questions are straightforward, common-sense, inexpensive-to-collect, and acceptable to users.

---

[44] **Schreiner *et al.***, 2014; **Zeller**, 2004.

[45] The statistical criterion is not the *p* value of an estimated coefficient 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 Bolivia. Customizing poverty-assessment tools by urban/rural does not improve targeting accuracy much.[46] Segment-specific tools may improve the accuracy of estimates of poverty rates,[47] but:

- They run a greater risk of overfitting[48]
- 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 Bolivia, the department of residence)[49]

[46] **Brown, Ravallion, and van de Walle**, 2018; **World Bank**, 2012; **Sharif**, 2009; Schreiner, **2006** and **2005**; **Narayan and Yoshida**, 2005; and **Grosh and Baker**, 1995.
[47] **Diamond *et al.***, 2016; **Tarozzi and Deaton**, 2009.
[48] **Haslett**, 2012.
[49] **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 that are 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 is it anything else).

To get cardinal units, a look-up table is used to convert scores to *poverty likelihoods*, that is, probabilities of having income below a poverty line. For the example of 100% of the national line, scores of 40−41 correspond with a poverty likelihood of 48.2 percent, and scores of 42−43 correspond with a poverty likelihood of 37.9 percent (**Figure 1**).

The poverty likelihood associated with a score varies by poverty line. For example, scores of 40−41 are associated with a likelihood of 48.2 percent for 100% of the national line but with a likelihood of 30.9 percent for the $5.50/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 income below a given poverty line.

For the example of 100% of the national line and a score of 40−41 (**Figure 22** below), there are 6,038 (normalized) households in the construction sample. Of these, 2,908 (normalized) have income below the poverty line. The estimated poverty likelihood associated with a score of 40−41 is then 48.2 percent, because 2,908 ÷ 6,038 ≈ 0.482 = 48.2 percent.

The same method is used to calibrate all scores with poverty likelihoods for all 15 supported poverty lines.[50]

[50] If needed to ensure that likelihoods never increase as scores increase, likelihoods across adjacent scores are averaged 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 (%)
0–25	3,036	÷	3,283	=	92.5
26–29	2,680	÷	3,193	=	83.9
30–32	2,560	÷	3,477	=	73.6
33–35	3,034	÷	4,533	=	66.9
36–37	2,631	÷	4,298	=	61.2
38–39	2,092	÷	4,343	=	48.2
40–41	2,908	÷	6,038	=	48.2
42–43	2,423	÷	6,401	=	37.9
44–45	2,296	÷	7,001	=	32.8
46–47	1,978	÷	7,365	=	26.8
48–49	1,696	÷	7,577	=	22.4
50–51	1,200	÷	6,650	=	18.0
52–53	729	÷	6,821	=	10.7
54–55	493	÷	6,588	=	7.5
56–57	347	÷	5,577	=	6.2
58–60	231	÷	6,520	=	3.5
61–65	131	÷	6,724	=	1.9
66–100	22	÷	3,611	=	0.6

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 income.[51] The fact that some choices in scorecard construction are informed by judgment in no way impugns the objectivity of the estimated likelihoods, as that 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 $2.718281828^{\beta X}$ x $(1 + 2.718281828^{\beta X})^{-1}$, 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 Bolivia's 2019 EH who have income 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.

[51] 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?

Estimation error is the distance and direction by which a scorecard's estimate tends to differ from the true value in the population.

For example, the estimation error of Bolivia's scorecard for estimates of head-count poverty rates in a single time period by 100% of the national poverty line is +2.4 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 81.5 percent, and if the estimation error is +2.4 percentage points, then an improved estimate is 81.5 − (+2.4) = 79.1 percent.

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

Estimation errors are reported for estimates of head-count poverty rates in a single time period for the 15 supported poverty lines for Bolivia.

The estimation errors are derived *out-of-sample*. This means that the scorecard (made from the construction sample from the 2019 EH, **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 income-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 (2019).

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.

The unknown degree and consequences of these inevitable violations of the scorecard's assumptions means that actual estimation errors will differ from those reported here in unknowable ways.[52] Still, the estimation errors (and margins of error) reported here are the best available, and it makes sense to account for them.

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.[53]

It is possible to compare estimated and observed poverty rates because the 2019 EH records actual (not estimated) income-based poverty status for households in the validation sample. The observed (not estimated) poverty likelihood in the 2019 EH 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 the 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 3.1**).

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

---

[52] Estimation errors due to being out-of-time can be measured with post-2019 data (say, from a future EH). Of course, future EH data is not yet available, and even after it is available, there will still be some unknown out-of-time error (and out-of-sample error will still be completely unknown).

[53] 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 Bolivia overall. The population's true value is taken as the value in the 2019 EH (even though the EH is itself only a sample).

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.[54] 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.[55]

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 Bolivia's 15 supported lines.

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.[56]

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

---

[54] 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.
[55] See **Schreiner** (2021) for details on the α factor and the formulas for estimation errors, margens of error, and ideal sample sizes in **Annex 5** and **Annex 6**.
[56] 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" (at least 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.

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.

To illustrate, suppose that the adjusted estimate of the head-count poverty rate for 100% of the national line is 79.1 percent and that the sample size is n = 1,024. Given 90-percent confidence,[57] the margin of error is ±3.5 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 79.1 − 3.5 = 75.6 percent to 79.1 + 3.5 = 82.6 percent, with the most-likely true value being the center of the range (the 79.1-percent estimate).

Said another way, "With 90-percent confidence, the estimate has a margin of error from 75.6 to 82.6 percent." This means that the true population value has a:

- 5-percent chance of being less than 75.6 percent
- 90-percent chance of being between 75.6 and 82.6 percent
- 5-percent chance of being greater than 82.6 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 Bolivia probably is indeed pro-poor if the scorecard estimate of the head-count 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. This is because 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-Bolivia rate for this line of 37.2 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 Bolivia overall (37.2 percent) and thus that the program may not actually be pro-poor.

---

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

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.

A5.2.3 Margins of error for estimates of poverty rates in one time period for the Bolivia 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 Bolivia scorecard are ±3.6 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 ±3.6 percentage points or less of the error-adjusted estimate.

A5.2.4 How to calculate margins of error

The **ProveIt™-brand reporting and analysis tool** calculates margins of error for all scorecard estimates discussed here. Analysts may also use the formulas below.

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:

$\pm c$ 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.

Suppose that the following are given:

- A confidence level that corresponds with z
- A sample-based estimate $\hat{p}$
- A population size N
- A sample n, and
- An adjustment factor α for a specific poverty line from **Figure 2**

Then the formula[58] 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, Bolivia's 2019 EH gives a direct-measure head-count poverty rate for 100% of the national line of $\hat{p}$ = 37.2 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.[59] Bolivia in 2019 had a population of households (not people) of N = 3,444,751, and the EH sample size was n = 11,839. Given a desired confidence level of 90 percent, z is 1.64. The margin of error $\pm c$ is then about ±0.7 percentage points:

$$\pm z \cdot \alpha \cdot \sqrt{\frac{\hat{p} \cdot (1-\hat{p})}{n}} \cdot \sqrt{\frac{N-n}{N-1}} = \pm 1.64 \cdot 1.00 \cdot \sqrt{\frac{0.372 \cdot (1-0.372)}{11{,}839}} \cdot \sqrt{\frac{3{,}444{,}751-11{,}839}{3{,}444{,}751-1}}.$$

This implies a 90-percent chance that Bolivia's true head-count poverty rate for 100% of the national line in 2019 is in the range from 37.2 − 0.7 = 36.5 percent to 37.2 + 0.7 = 37.9 percent.

[58] 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.
[59] For scorecard estimates, α for a given poverty line is found in **Figure 2**.

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 head-count poverty-rate estimate.

To illustrate, the baseline example in **Section** **3.1.1** has an estimated poverty rate of 79.1 percent. With 70-percent confidence, the margin of error is about ±40.9 percentage points,[60] or from 79.1 − 40.9 = 38.2 percent to 79.1 + 40.9 = 120.0 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.[61]

The estimated number of people in participating households in the example in **Section** 3.1.2 is 6,000,[62] so the lower limit of the 70-percent margin of error for the estimated number of poor people is 6,000 · 0.382 = 2,292. The upper limit is 6,000 · 1.000 = 6,000. This example estimate—based as it is on a sample of two households—is better understood not as "all people in participating households are poor" but rather as "at least about 40.9 percent are poor".

---

[60] The example in Section **3.1** has an estimate of 79.1 percent with N = 1,000, n = 2, and α = 1.37 (**Figure 2**). For 70-percent confidence, z = 1.04. The margin of error ±c for the head-count poverty-rate estimate is then

$$\pm 0.409 \approx \pm 1.04 \cdot 1.37 \cdot \sqrt{\frac{0.791 \cdot (1 - 0.791)}{2}} \cdot \sqrt{\frac{1{,}000 - 2}{1{,}000 - 1}}.$$

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

[62] 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.

A5.2.7 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:

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

- z, α, N, and n are defined as above
- $\hat{p}_{up}$ is the estimated share of members of sampled households that rise above the poverty line from below
- $\hat{p}_{down}$ is the estimated 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

Illustrating with the earlier example of one sample scored twice (**Section 3.2.1**), $\hat{p}_{up}$ 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 | −0.473 + −0.502 | = +0.975), divided by the sum across all sampled households of each household's average household size across baseline and follow-up of 5.5 + 7.5 = 13.0 (from rows 3 and 4, column G). Thus, $\hat{p}_{up}$ = +0.975 ÷ 13.0 = 0.075.

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 (5.5 + 7.5 = 13.0) gives $\hat{p}_{down}$ = 0.000 ÷ 13.0 ≈ 0.000.[63]

The household-size-weighted average of the number of years between interviews y is 3.00 (from row 9, column M in **Figure 11)**.

---

[63] $\hat{p}_{down} - \hat{p}_{up}$ is the estimated net poverty-rate change. In this example, $\hat{p}_{down}$ ≈ 0.000 and $\hat{p}_{up}$ =. 0.075, so 0.000 − 0.075 = −0.075, which is the estimated −7.5 percentage-point decrease (improvement) in the poverty rate 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,[64] the margin of error ±c is about ±0.074 ≈

$$\pm\frac{1.04\cdot 1.14}{3.00}\cdot\sqrt{\frac{0.075\cdot(1-0.075)+0.000\cdot(1-0.000)+2\cdot 0.075\cdot 0.000}{2}}\cdot\sqrt{\frac{1{,}000-2}{1{,}000-1}}.$$

The example's estimated net annual poverty-rate change is about −2.5 percentage points (**Figure 11**), so the 70-percent margin of error is from −2.5 − 7.4 = −9.9 to −2.5 + 7.4 = +4.9 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 be strongly negative, close to zero, or strongly positive.

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 2.5 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 head-count poverty rate.

To illustrate with the example in **Section 3.2.1** for one sample scored twice, the estimated annual net change in the poverty rate is about −2.5 percentage points. As just shown, the tiny sample size of n = 2 means that the 70-percent margin of error runs from −9.9 to +4.9 percentage points.

[64] **Schreiner**, 2021.

The estimated average number of on-going participating people per year is 5,450 (**Figure 11**).[65] Thus, the lower limit of the 70-percent margin of error for the estimated annual net change in the number of poor people is
5,450 · (−0.099) ≈ −540 (a net decrease in poor people), and the upper limit is
5,450 · (+0.049) ≈ +267 (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 a true change of zero.

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. There are *n* households sampled and interviewed at baseline, and another *n* households sampled and interviewed at follow-up.

Illustrating with the example for two independent samples in **Section 3.2.4**:

- *z* = 1.04, assuming a desired confidence level of 70 percent
- α = 1.10, the adjustment factor (regardless of poverty line) for this estimator[66]
- *y* = 2.73, the years between the average interview at baseline and follow-up
- $\hat{p}$ = 0.815, 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 ±0.163 ≈ $\pm\frac{1.04\cdot1.10}{2.73}\cdot\sqrt{\frac{2\cdot0.815\cdot(1-0.815)}{2}}\cdot\sqrt{\frac{850-2}{850-1}}$.

The example's estimated net annual poverty-rate change is −2.7 percentage points (**Figure 12**). Thus, the 70-percent margin of error is
from −2.7 − 16.3 = −19.0 percentage points
to −2.7 + 16.3 = +13.6 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.

---

[65] See footnote **62**.
[66] **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 head-count poverty rate.

To illustrate, the example in **Section 3.2.4** for two independent samples estimates the annual net change in the poverty rate as −2.7 percentage points. As just shown, the 70-percent margin of error runs from −19.0 to +13.6 percentage points.

The estimated average number of on-going participating people is 4,225.[67] 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 $4{,}225 \cdot (-0.190) \approx -803$ (a net decrease in poor people), and the upper limit is $4{,}225 \cdot (+0.136) \approx +575$ (a net increase in poor people). The margin of error again shows that the estimate does not reveal much about the true value in the population.

[67] See footnote **62**.

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[68]
- 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.

[68] Academic conventions for levels of confidence, when 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.[69]

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 = 1.37 (**Figure 2**)
- The pre-estimation expected poverty rate is the all-Bolivia rate for 100% of the national line in 2019, so $\tilde{p}$ = 37.2 percent = 0.372 (**Figure 10**)
- The desired margin of error $\pm c$ = ±3.0 percentage points = ±0.030

Given these hypothetical values,

$$n = 10{,}000 \cdot \left(\frac{1.28^2 \cdot 1.37^2 \cdot 0.372 \cdot (1-0.372)}{1.28^2 \cdot 1.37^2 \cdot 0.372 \cdot (1-0.372) + 0.03^2 \cdot (10{,}000-1)} \right) \approx 740.$$

[69] 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:

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

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
- 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-Bolivia rate for 100% of the national line: $p_{\text{pre-baseline}} = 37.2$ percent = 0.372 (**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 $2 \cdot \left(\frac{1.28 \cdot 1.14}{0.03} \right)^2 \cdot [-0.01 + 0.016 \cdot 3 + 0.56 \cdot 0.372 \cdot (1 - 0.372)] \cdot 1 \approx 799.$

The follow-up sample size is also 799.

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 $n = 2 \cdot N \cdot \left(\frac{z^2 \cdot \alpha^2 \cdot \tilde{p} \cdot (1-\tilde{p})}{z^2 \cdot \alpha^2 \cdot \tilde{p} \cdot (1-\tilde{p}) + c^2 \cdot (N-1)} \right)$.[70]

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

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

$$2 \cdot 10{,}000 \cdot \left(\frac{1.28^2 \cdot 1.10^2 \cdot 0.372 \cdot (1-0.372)}{1.28^2 \cdot 1.10^2 \cdot 0.372 \cdot (1-0.372) + 0.03^2 \cdot (10{,}000-1)} \right) \approx 979.$$

The sample size at follow-up is also $n = 979$.

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

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**.

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**.

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**.

Instituto Nacional de Estadística. (2020a) “Encuesta de Hogares: 2019”, **link**.

_____. (2020b) “Actualización de la Canasta Básica de Alimentos y de las Líneas de Pobreza”, **link**.

_____. (2020c). “Cálculo de Líneas de Pobreza: Documento Metodológico”, **link**.

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**.

_____. (2015a) “Scorocs Simple Poverty Scorecard Tool: Bolivia”, **link**.

_____. (2015b) “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**.

_____. (2009) “Scorocs Simple Poverty Scorecard Tool: Bolivia”, **link**.

_____. (2007) “Scorocs Simple Poverty Scorecard Tool: Bolivia”, **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, **link**.

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**.

Zeller, Manfred. (2004) "Review of Poverty Assessment Tools", **link**.