

Data collection

A practical guide to collecting data in
developing countries

Vojtěch Bartoš

Ian Levely

About the guide

This guide explains the basics of data collection. It discusses common **errors and biases** that may arise during sample selection, how questions can produce misleading information when asked improperly, and **how to overcome these issues**. It also presents sample questions on key topics: **wealth, health, education and hygiene and sanitation**. After reading this guide, you should be able to **design a data collection** instrument from the **selection of respondents** down to **data entry**.

About the authors

Vojtěch Bartoš, CERGE-EI, Prague, bartos.vojta@gmail.com.

Ian Levely Institute of Economic Studies at the Charles University, Prague, ianlevely@gmail.com.

Both authors are members of the Group for Analysis of Poverty and Inequality (GAPI).
www.gapiresearch.org

©2014 Člověk v tísni o.p.s., Šafaříkova 635/24, 120 00 Praha 2, Česká republika. All rights reserved.

Phone:(+420) 226 200 400

Email:mail@clovekv tisni.cz

Web:www.peopleinneed.cz

Table of contents

<i>About the authors</i>	2
<i>Table of contents</i>	3
Chapter 1: Choosing who to survey	4
<i>Bias in sampling</i>	4
<i>Sample Selection</i>	7
Multi-stage cluster sampling	8
Weighted sampling	9
<i>Non-recommended methods of sample selection</i>	9
Convenience samples	9
Snowball sampling	10
<i>Surveys over time</i>	10
<i>Choosing the sample size</i>	11
<i>Key reading</i>	11
Chapter 2: Designing Surveys	12
<i>Error and bias in responses</i>	12
<i>Accuracy of responses</i>	13
<i>Survey length</i>	14
<i>The process of creating surveys</i>	15
<i>Key reading</i>	16
Chapter 3: Measuring Outcomes	17
<i>Measures of Wealth</i>	17
Income	17
Consumption	20
Assets	22
<i>Debt</i>	22
<i>Health</i>	23
<i>Infrastructure and sanitation</i>	24
<i>Education</i>	25
<i>Choosing survey questions</i>	26
<i>Key reading</i>	26

Chapter 1: Choosing who to survey

- The best way of choosing who to survey is to select randomly from the population.
- The method by which individuals or households are selected for surveys can have a large influence on the results, possibly introducing a bias.
- Using multi-level cluster sampling makes it possible to access random samples of a population in the absence of a comprehensive list of the population.
- Weighted samples allow you to ensure that you get enough observations of each type that you need to make meaningful comparisons.
- Samples of convenience should be treated with care and are only useful in specific contexts.

In most cases, there won't be enough resources available to measure the outcomes of every program participant. Likewise, there will likely be individuals fail to complete surveys or are unreachable for various reasons. In Bartos and Levely (2014): "Impact evaluation" we discuss how evaluations should be designed in a way that allows for a proper comparison group to assess the effects of program interventions. Just as selection of individuals into the *treatment* can depend on personal characteristics that may also affect outcomes, the likelihood that someone ends up in the survey can depend on similar factors.

This is a major issue in performing baseline or needs assessment surveys. If you want to get an idea of the characteristics of the average household or individual living in an area, the selection of the sample of that population that is chosen for the survey will have a very large effect on the outcome.

In this chapter, we first explain how selection of survey participants can change results, and then we continue by explaining a few methods that can be used for sampling.

Bias in sampling

A common belief is that some data is better than no data, so even if you don't manage to reach every individual that you intend to, or even if there are some flaws in the way that data is collected, the results of such an imperfect study still may tell you something that you would not otherwise know. While it's certainly true that all surveys will fall short of perfection, this line of thinking can lead to misleading conclusions and counterproductive policies. In other words, **poorly collected data can sometimes be worse than nothing at all!**

Let's say we want to survey people living in a village in a rural area of a developing country with a population of one hundred households, of which there are resources to reach twenty. Since we want the households we survey to accurately represent the village, we *randomly* select the twenty households from a list of all the households in the

village. Next, we send a survey enumerator to each household while we wait patiently in the mission office. When the enumerator comes back at the end of the day, she tells you that at five out of the twenty households, no one was at home and therefore the survey could not be completed. How serious of a problem is this?

To answer this question, we need to consider not only the *quantity* of data that we have lost, but the *quality* of the data that we have succeeded in collecting:

Scenario 1—Random Error: It could be the case that every single household had a similar chance of being empty when the enumerator arrived. In other words, the households were selected randomly by us, then randomly again by chance, according to who happened to be at home. This is not such a serious problem and only affects the cost of collecting data. The next day we could just choose more households. If the enumerator goes to a house and no one is at home, she can continue to the next on the list and will eventually collect data from twenty households.

Scenario 2—Selection bias: Alternatively, there could be one or more reasons that no one is at home in certain households. An extremely obvious example would be if the village has both Muslim and Christian residents and we collect data on a Friday or Sunday and half of the people are at mosque or church, respectively. A more subtle (and realistically overlooked) problem would be that households with younger, more economically active members would be more likely to be away during the day when the survey is conducted. In contrast to random error, this biases the sample selection process and our results no longer accurately represent the village's population. Our data will likely make the village look older, poorer and less economically active than it actually is, since we systematically select people with these characteristics by only interviewing those who are at home.

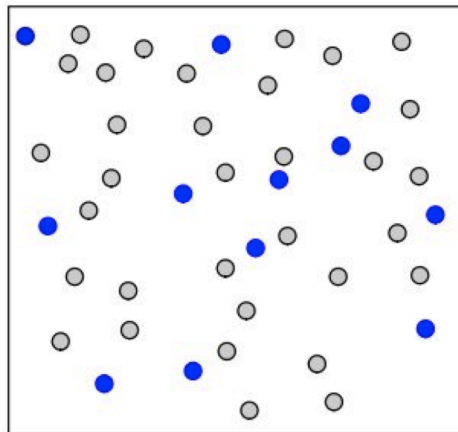
In this situation, the bias could be minimized by investing more resources towards increasing the response rate, perhaps by sending enumerators to the village later in the day when more people are at home or by having them return more than once to each village and track down the missing households. In the end though, we will most likely have to live with a certain degree of bias.

There are two main takeaways here:

- Firstly, we should **be aware of any potential biases and interpret results accordingly**. Any conclusions drawn from this data can only be applied to the 75% of the population that responded to the survey, and we can make an educated guess—based on our knowledge of the survey process and the population—about how the missing households might differ (are they likely younger, richer or poorer etc...).
- Secondly, the process by which the **sample is selected should be as unbiased as possible** and we should design the sampling method in such a way that we don't introduce any new, *unnecessary* sources of selection bias.

In this example, we assumed that there was a list of households in the village that we could use to randomly select households to visit. If you are interviewing clients of a particular program, you will likely have such a list, which makes the process easy: just use a method to choose a given number of these individuals *randomly* from the list.¹ Then, simply stick to the randomly selected clients as closely as possible and avoid any temptations to deviate from this randomization.

Figure 1.1: Random sample



Source: Authors

Sample selection during evaluation of self-help groups in Ethiopia

When one of the authors was evaluating self-help groups in Ethiopia, we first looked at list of all groups working with PIN (the "population"). Based on the resources available, we selected a small number of groups to interview from that list, using a random number table and set up dates to interview members of those groups for the study.

One of the randomly selected groups to be interviewed did not show up at the appointed time because of a communication problem between the partner NGO and the group. There was, however, another group that was close by and one of the partner's employees suggested that we interview those women instead. This was a tempting suggestion since the enumerators were idly standing by, not collecting data. However, doing this would bias the sample.

To see how, we'll consider a couple of explanations for why we would have interviewed the second group in place of the first:

1. Why didn't the randomly selected group get the message? It could be a simple error on the part of the employee who was asked to contact them. Possibly, it had nothing to do with the group itself. But what if she would have been less likely to contact

¹ An easy method is to generate random numbers in an excel table. In a pinch, you can write numbers on slips of paper and draw them from a hat.

certain groups, such as those with members that were in closer contact with the NGO or those geographically closer to the office? In an extreme case, the partner NGO could have been trying to keep us from talking to a particular group of clients who they feared would reflect poorly on their organization—either intentionally or unconsciously.

2. Secondly, why was group at the office available? Did all of the groups have the same chances of being selected in this haphazard way? Maybe the NGO employee would only suggest interviewing groups that reflect positively on their organization (even if it is not a conscious decision). The group that happened to be at the office was there for a regular meeting, which meant that they meet during the week and hold these meetings at the main office, whereas other self-help groups have meetings on weekends at smaller, regional offices.

Keep in mind: By relenting and interviewing the most conveniently reachable clients, we systematically choose a certain type of person and the quality of our data decreases, potentially leading to incorrect conclusions.

Sample Selection

Thus far, we've considered situations in which we select a group of people to survey from a longer list of individual or households that we know about and can easily contact. This is a best-case scenario, however, and especially for needs-assessment surveys there probably will not be a list of village residents or potential clients that you can randomly draw from.² In this case, we need a method for selecting interviewees from the community as randomly as possible.

Give clear instructions on how to choose interviewees. An extremely bad approach would be to send an untrained enumerator to a neighborhood and instruct her to interview 100 people. Her main motivation is to do the job as quickly as possible by choosing the people who are easiest to find. These people might be more likely to live near the center of the town. Some people may be hesitant to complete surveys, perhaps because they are busy or distrustful. People who are employed or running a business are more likely to be busy during the day than those without work. All of this introduces bias. Luckily, we can eliminate a great deal of this by carefully designing a clear procedure for selecting people to interview, or a *sampling frame*.

Choose a population of interest. But first, we need to define the *population of interest*, or those people that we are interested in collecting data from. If we are conducting a needs assessment survey, the population of interest might be everyone living in the area or perhaps only everyone in the area who meet some set of criteria for the project (i.e. the poorest individuals or households, women with children, etc...). If we are collecting data

² Even if such a list does exist, some households might be missing from the list. These households could be newer, further away from the center or socially marginalized families. Excluding them from the survey could introduce a significant bias..

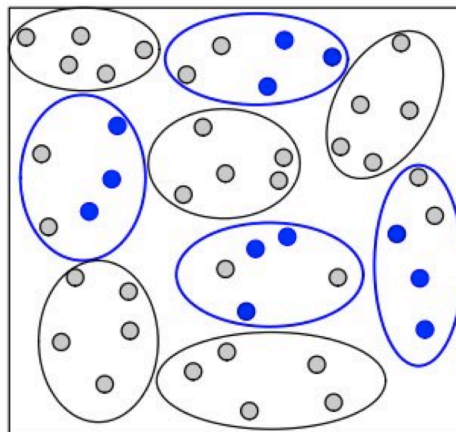
to measure the impact of a program, then our population of interest will be the control and treatment groups (see Bartos and Lively (2014) "Impact evaluation", chapter 1).

Above, we considered a survey in a village where we had a list of all the households. But what if we do not have such a list? We could make our own list of households and then randomly select households from that list, but this would be time consuming and probably not feasible.

Multi-stage cluster sampling

An alternative is *multi-stage cluster sampling*, which involves randomizing at two or more levels. For example, again taking a village as population, you might divide the area into geographic units of roughly the same population, or "clusters," then select a few of the clusters at random. Then, individual respondents are selected, again randomly, from within each cluster and surveyed. Since each individual has an equal chance of being selected, cluster sampling leads to unbiased results. However, since observations are more concentrated, there is less variation, which means that results are less precise. (See figure 1.2)

Figure 1.2: Clustered sample



Source: Authors

Let's say that you are working in an urban area and are interested only in poorer households (maybe below some pre-determined threshold) and living in certain neighborhoods, defined by the scope of the project. This is the population of interest and your task is to construct a sampling procedure that is as representative as possible of all the people who meet this criteria. To apply two-stage cluster sampling, you would first use a map of the neighborhoods and divide it into a grid. Then you could randomly select a number of geographic points on this grid and then provide clear instructions to enumerators to choose a house close to that point (for example the first house to the South, for example). Hand-held GPS devices make this process much easier.

Weighted sampling

Even if sampling is random, it is possible to end up with no observations from one segment of the population and more than enough observations from another group, just by chance. Especially if you have information about the characteristics of the population at large, for example from a Demographic and Health Survey³ or national census, you might employ a *weighted sampling* technique. This involves targeting certain types of individuals—say those living in a particular neighborhood—and deciding that you will conduct a certain number of interviews in that area. Later, when you analyze the data, you take into account how the sample was weighted when drawing conclusions about the data.

One possibility is to weight the sample by neighborhood to insure that you have enough observations from each area to draw meaningful conclusions. So, rather than randomly selecting individuals from the wider pool, you determine a number of observations that you will collect from each neighborhood. Later, when you conduct the analysis, you should consider the number of people who live in each neighborhood if you make any general conclusions about poor people living in the city.

Non-recommended methods of sample selection

There are a couple additional methods of sampling that you may come across that should be approached with extreme caution.

Convenience samples

Convenience samples are groups of easily accessible individuals who might be used to collect survey data. In our village of one hundred households, one option might be to use the village school as a way of identifying households. While this is practical and saves resources, there is a major disadvantage, in that you will only identify those households with children enrolled in school. There will be some households without school-age children and some households who do have school-age children but who are not enrolled in school that will be left out, leading to an obvious bias.

This may be an appropriate methodology if we are, in fact, only interested in families with children enrolled in school—i.e. if this is our population of interest. But, if our population is the entire village, this convenience sample leaves out an important group of households.

Some uses of convenience samples can lead to useless and misleading conclusions. Surveying clients of a rural health clinic in a developing country on health-related issues, for example, would tell you very little about health-care needs and coverage for area as a whole, since you wouldn't collect data on the population that doesn't frequent clinics. This may be due to better health, or because they can't afford or are less willing to pay for healthcare.

³ <http://www.measuredhs.com/>

Snowball sampling

A commonly used technique to access hard-to-reach populations (such as ex-combatants, AIDS patients or illegal drug users) is called *snowball sampling*. This method involves finding a small group of people who meet the narrow criteria of the survey, then using those individuals to find more respondents. Each "wave" of respondents is used to contact the next, so the sample slowly grows like a snowball rolling down a hill. This is an extremely sloppy and haphazard way of sampling, as **each round of the survey introduces a bias**, since respondents are most likely to refer you to those contacts who they believe are likely to attend the survey. Moreover, you will systematically under-sample individuals who have smaller social networks. While this method is used by some researchers in practice, we do not recommend it.

Surveys over time

Much more can be learned from surveys that measure outcomes from the same individuals at several points in time than can be learned from collecting a "snapshot" of data at one point in time. For example, if we have a randomized evaluation, we definitely want to collect data both before and after the program from both clients and non-clients. Although this can be logistically difficult, especially with the non-clients, it is worthwhile and improves the value of the study greatly.

If there are fewer resources for surveying in further rounds of the study, the best plan of action is probably to randomly select a smaller sample from the larger original respondents. You might, alternatively, decide to concentrate on part of the sample that is of particular interest. There is nothing wrong with this, as long as you recognize that the results are limited.

In Bartos and Lively (2014) "Impact evaluation, chapter 1 we discuss the problem of *attrition* in designing impact evaluations: some people will drop out of the program, and these people likely have different characteristics from those who stay. A related problem involves finding survey respondents over time. If you survey 20 households in a village of 100, then go back a year later and try to collect data from the same individuals, you might not succeed in finding all of the same households. As we discussed earlier, this may be due simply to random error, in which case we do not have a problem. However, if the characteristics of the missing individuals cause some households to be more likely than others to be represented in the second (or following) rounds of a survey, then attrition will bias survey results. Migration — for example to cities for work — is one common culprit.

Again, the problem does not make the data completely useless, but limits its generalizability to those households that you remain in the study.

Choosing the sample size

A common question is how many surveys need to be collected. Unfortunately, there is not simple answer to this question. There are two guidelines that help you to determine how large of a sample you need:

1. Increasing the sample size allows statistical identification of smaller effects.
2. The more varied the data, the larger the sample size that you'll need.

When running an impact evaluation the primary concern is to identify the effect of the program. The smaller the sample size, the *larger* the effect of the program will have to be in order for the survey to identify a statistically significant effect. See the online appendix for more information on how to determine the appropriate sample size for the study.⁴

Key reading

For more detailed information about sample selection refer to the World Bank's Haughton, Khandker (2009). Handbook on Poverty and Inequality. The World Bank, Washington DC. Available online at: <http://hdl.handle.net/10986/11985>

⁴Available online:
<https://www.dropbox.com/s/2wsd0zofxbx3apo/power%20calculation%20tool.xlsx>

Chapter 2: Designing Surveys

- The way that questions are asked can introduce bias into the results.
- The accuracy of responses to survey questions is influenced by people's memories, by the enumerators asking the questions and by beliefs about how answers might influence the benefits respondents receive from the NGO.
- The accuracy of responses to survey questions decreases with the length of the survey.
- Creating surveys should be a dynamic process that includes discussions with knowledgeable locals, interviews and testing.

Error and bias in responses

In the previous chapter, we discussed how choosing the appropriate sample can bias results. Here, we turn towards how bias can arise in the way that questions are asked and answered.

Random error. If we ask about cash income, for example, it is unlikely that many people will know this figure precisely, but will rather give an approximate answer. If we ask a group of people with an average income of \$100 how much money they have made, and each person gives an answer that is either \$10 higher or \$10 lower than the true amount with an equal probability, the data is less precise, but if we interview enough people, the average will still be \$100. This is *random error*.

Random error can make it more difficult to find patterns in data. If some of the participants had participated in a livelihood development program that raised incomes, the random error in reporting income might hide the effect of the program. Random error can lead to *false negatives* (also known as type-II errors), or situations in which there is a significant effect, but the results do not demonstrate this. However, as the number of individuals surveyed increases, the chances of a false negative decrease: the larger the number of observations, the less random error matters.

Systematic bias. In the previous example, imagine that instead of randomly reporting income \$10 higher or lower than the actual value, all individuals are likely to forget about at least some income. Now *everyone* reports a figure lower than their true income, and the average income measured by the survey will be lower than the true average of \$100. Since everyone reports a number lower than the true figure, surveying more individuals won't make statistical tests more accurate. In other words, more observations does not eliminate the effect of *systematic bias*, in contrast to *random error*.

Now, imagine that you collect data on annual income and that half of those surveyed have taken part in a livelihood development project, and part of this project involved education on personal finance. The project might increase the chances that participants will report all of their income, making non-participants comparatively more likely to forget some income and thus report lower incomes. If we observe that program

participants have higher incomes on average, there are (at least) two possible reasons for this: the program actually raised incomes, or the program made participants more likely to *remember* their income better and thus *report* higher incomes even if there was no change. Of course, both of these effects could occur simultaneously, but the point is that we don't know how much of the difference is due to the change in income and how much is due to more accurate reporting. Thus, results of the survey will make the program appear more effective than it actually was. This illustrates how *systematic bias* can lead to serious problems in data analysis.

The question is then how to avoid such bias. There is no easy answer to this question, since the problems in collecting accurate data will vary according to the population being studied. Generally, it is important to try to anticipate such bias and design survey questions in a way that minimizes the potential for misleading results. For example, one might ask not only about annual income, but also about weekly or monthly income, which may be more accurately reported (although this can produce a bias of its own, see Figure 3.1 in Chapter 3).

But, regardless of how well a study is designed and how carefully data is collected, it is unfortunately practically impossible to collect data that is 100% accurate. This is true for a variety of reasons: survey respondents themselves may not have accurate information about the issues you are investigating or have faulty memories. If you ask individuals about household characteristics, they may not have precise answers because they only have a rough idea about family members conditions or behaviors.

Accuracy of responses

One question present in all survey work is "how much can you trust the information that people provide you with?" since the information that you are after is filtered through the respondents.⁵

- *Recall error*: People tend to under-report small things because some instances slip their memory. For example, asking the question: "how many cups of tea have you had in the past week" will likely produce a result lower than the true value, since people mentally tally the number by trying to remember each cup.

→ The only solution to this problem is to collect more detailed information when possible and to ask in a way that helps subjects remember their past actions, and to realize that data of this type is likely biased downwards.

- *Telescoping*: By contrast, people tend to *over* report the number of large purchases that they make in a give time period.⁶ If we ask if you have purchased a mobile phone within the last year, and you have done so thirteen months prior,

⁵ For a similar list, see Deaton, Angus and Margaret Grosh. 2000. "Consumption," in *Designing Household Survey Questionnaires for Developing Countries: lessons from 15 years of the Living Standards Measurement Survey, Vol 2*, The International Bank for Reconstruction and Development, The World Bank.

⁶ UN Department of Economic and Social Affairs, Statistics Division, "Household Sample Surveys in Developing and Transition Countries,"

you might remember the purchase but not exactly when it occurred and erroneously report that you have purchased the item "within the past year."

→ One way to deal with this issue is to use *fixed reference points* (e.g. "since Christmas" instead of vague time frames (e.g. "over the past year").

- *Prestige effects*: reporting assets, consumption or income that is *higher* than the actual level in order to appear better off than—or embarrassment of low consumption or social problems.

→ Proper survey introductions and explaining that the results will not be shared and questions that are as specific as possible help to mitigate this problem.

- *Surveyor effect*: The person asking the questions will invariably have an effect on the way the questions are answered, and even when comparing data collected by two experienced enumerators there might be bias introduced.

→ The first defense against this effect is to make sure that enumerators are asking questions in a uniform way. It is advisable to go over the questionnaires with all data collection staff to agree on standard ways of asking questions and recording answers.

→ Secondly, carefully randomize which respondents conduct the interview with whom. If, for example, you assign one enumerator to interview everyone in the treatment group and a second to interview everyone in the control, surveyor effects might cause differences in the outcome and it will be impossible to separate this from the treatment effect. If you randomly assign individuals from each group to one of the two enumerators, then this problem is much less severe.

- *NGO effect*: As you are representing an NGO when asking information, survey respondents may (erroneously) believe that they will receive more benefits if they make themselves appear to be in more need and thus report *lower* incomes or consumption than the real.

→ Carefully explain in the introduction to the survey that the information will be kept private and will not be used to determine the individual dispersal of benefits (assuming this is true).

→ Add crosscheck questions that will allow you to identify inconsistencies that might indicate responses aren't accurate.

Survey length

Many of the recommendations that we've made here for increasing the accuracy of collected data involve more or more detailed survey questions. Unfortunately, real

estate on questionnaires is a precious commodity. Increasing the length of surveys means that each interview takes up more time, which makes each observation more expensive. In practice, with limited time and money to collect data, we face a trade-off between the number of surveys that can be administered and number of questions on each individual survey.

Additionally, the accuracy of data provided by respondents tends to decrease as the interview progresses and respondents become tired and less interested in the process. (This problem is referred to as *survey fatigue*).

This means that designing a survey involves a necessary tradeoff between quantity and quality of data, and what goes on the final version of the survey is often a very difficult decision to make. Running pilot sessions (see below) to decide which questions are most informative can help in this process. If almost everyone gives the same response to a question in a pilot session, you might want to consider cutting that question, as it is unlikely to end up telling you anything meaningful.

The order of questions can also affect the answers given. Asking about sensitive material might make people nervous and uncomfortable when answering the remaining questions. It might be better to put those questions at the end of surveys. Of course, since the sensitive material might be what you are most interested in, which might justify moving the question towards the beginning. Just be aware of the tradeoff and use your best judgment.

The process of creating surveys

Creating a survey is a dynamic process, and by the time the survey is finally implemented, you should have gone through a dozen or so versions. The process starts by considering the information that you need to collect in order to answer the question posed by the study.

1. **Selecting questions.** Questions should then be drafted. You can turn to examples of previous surveys for an idea of how to phrase questions and which options to provide for multiple-choice questions (see Chapter 3). It cannot be stressed enough that the cultural, social and economic context affects how and what questions should be asked. Given this, it is advisable to consult surveys run *in the particular country* for inspiration.
2. **Find other questions when talking to local people.** Qualitative interviews or focus-group discussions can be useful in this process. You might find that people voice specific concerns or raise issues that hadn't occurred to you. Based on

anecdotes alone this information is hard to apply, but it can be very useful in forming the basis for survey questions.

3. **Mind possible cultural issues.** After you have a draft of the questionnaire prepared, consult with local staff to discuss any cultural issues you may have overlooked.
4. **Pilot your questionnaire.** Next, you should run a small-scale pilot to put the questionnaire to the test. Initially, don't count on using this data in the final results, but choose participants similar to those who will be in the study and interview them to see how they react to questions. Look at the data for questions that produce confusing results. If everyone has the same answer to a particular question, consider cutting it.
5. **Identify confusions or difficult questions.** After running the pilot, you should consult with the enumerators to identify any confusing or difficult questions. Identifying these problems from the beginning will make your life much easier later on, and it is well-worth investing time into a proper pilot.

Key reading

Deaton, Grosh (2000). Designing Household Survey Questionnaires for Developing Countries: Lessons from 15 years of the Living Standards Measurement Survey. The International Bank for Reconstruction and Development, The World Bank, Washington DC.

Chapter 3: Measuring Outcomes

- Income, assets and consumption are closely related yet distinct measurements of well-being. Knowing which measures to use in which context allows for a better understanding of an individual or household's economic situation.
- Measuring abstract aspects of human capital such as health and education requires collecting information on proxies that can be observed and measured. The choice of which information is collected will impact the results of the analysis.
- Well-being is extremely hard to quantify and is not necessarily captured by income and assets alone.

The goal of livelihood interventions is to increase productive capital, with the ultimate goal of increasing quality of life.⁷ Since these are somewhat abstract concepts, measuring these outcomes can be challenging, and often involves using indirect proxies. In this chapter, we consider various indicators that measure wealth, health, access infrastructure and education.

Measures of Wealth

There are three major categories of measures of wealth that we might can measure: **income**, **consumption** and **assets**. All three of these measures are closely related and each has its attributes and downsides as a measure of well-being. Ideally, a well-designed survey will include questions on all three categories and the combined results paint a complete picture.

The level of debt that a household carries also affects their economic situation. Consumption in the present time might be financed by loans that will have to be paid back in the future.

The choice over which questions you choose to include on a survey ultimately depends on the question the study seeks to answer and what results you expect to see. There is no easy answer as to which of these measures is better overall. In the following we discuss the pros and cons of various measures of well-being in each category.

Income

In developed countries asking about income alone gives a fairly accurate description of standard of living, as most people's income comes solely from formal sector work. The situation in developing countries is quite different. Firstly, when income is earned in the informal sector there are less records available and it is much more difficult to collect reliable information. Secondly, many transactions might not involve cash. Households may produce their own goods for consumption—especially production of food in rural areas—or may barter for goods or services without cash ever changing hands. This counts towards consumption and assets but not towards income—or at least not *cash*

⁷ See Chapter 1 in Bartos, Levely (2014): Livelihood development: A practical guide to administration of PIN livelihood development programs.

income. A third problem with measures of income, particularly at the household level, is that it does not directly measure well-being and might be misleading if spending is not equal between household members.

As unreliable as it may be, cash income is still a good starting point, and it's almost always worthwhile to gather at least some data on income. The level of detail questions on income should be designed carefully *according to the research question*:

Asking about cash earned: You can expect record keeping on cash income to be spotty at best, and memory of earnings—especially from micro-businesses—to be even less reliable. For this reason, it is advisable to elicit income based on each IGA in which an individual (or household) is engaged. The questionnaire could start off by asking about what types of work were completed over the previous month, then continue to ask about how much money was earned from each activity. This helps interviewees to remember what they have earned and also provides a more detailed report of economic activity, which might be interesting in and of itself if you are interested in knowing more about what types of IGAs people are involved.

Household vs. individual income: Household income is arguably a better measure of standard of living than individual income in many cases since household income is distributed among family members. However, there is a tradeoff here with reliability: obviously individuals will know more about their own economic activity than about that of other household members. If understanding household income is important to the study, you might consider asking about income from each household member or asking about each household income generating activity. It is important to bear in mind that household income is not always distributed equally between members, and this means that income gives a skewed result of well-being in some cases—especially with regards to women and children.

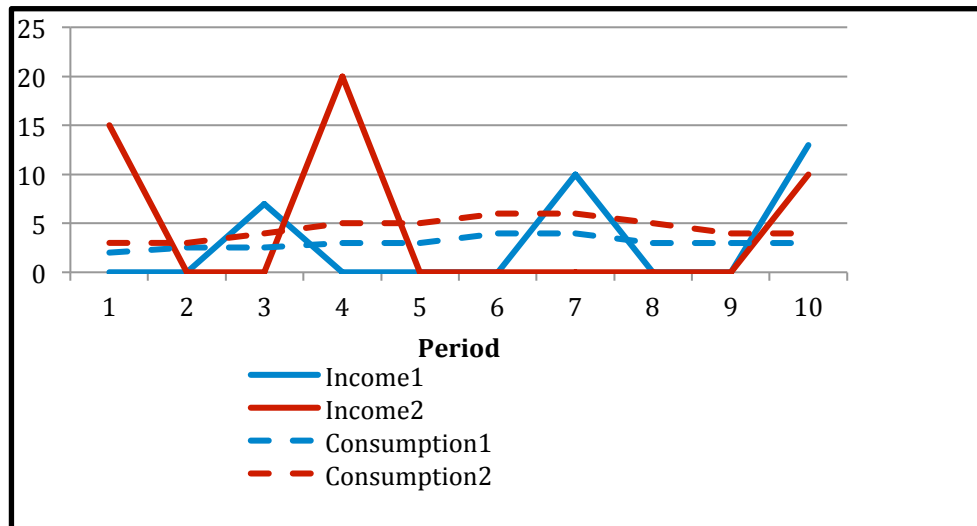
Time period: Income is measured over a given period of time. Asking about income over the past year is usually unrealistically ambitious and won't produce meaningful results. A month or week is more reasonable.

Stability of income: Income can vary greatly over time. Households may not receive a steady flow of cash on a weekly or monthly basis, but rather might earn nothing for, say, two months and then receive a large lump-sum of cash from a single transaction in the following month. This means that the narrower the time frame you ask about, the more variation might be expected. For example, when looking at cash income among farmers in a rural area, the main source of cash for the year may come from selling livestock. If you ask about income during the week or month that the sale is made, you'll observe a high amount of cash income, but nothing during the rest of the year. We would expect to see a similar effect with income from the sale of crops, which will vary greatly according to the season. The lesson here is to know the population well and design surveys in a way that accounts for problems like this and interpret results accordingly.

An important source of income for many poor households are remittances from relatives or friends living abroad or in large cities. Accounting for this income is important in putting together the financial situation of the household. Asking about

money received from relatives and friends within a specific time period has the same pitfalls as error in reporting income, but since the money usually is received in lump sums from money transfer services, such as Western Union, or mobile money transfers, people are likely to have a better idea of how much they have received and when.

Figure 3.1: Measuring income or consumption?



Source: Authors

Caption: In this diagram, income for two imaginary people varies over time, but consumption is smoothed (through saving and borrowing), even though overall consumption is equal to overall income across the 10 periods. If we measure income in period 4, we would conclude that person 1 has no income, while person 2 has very high income, but would conclude the exact opposite if we collected data in period 7. However, the consumption data gives us a much clearer picture of both individuals' standard of living, regardless of when we collect data.

- How much money have you earned in the past 7 days?
- How much money have you earned in the past month?
- Which of the following activities have you done over the past month?
 - Working your own land?
 - Working on someone else's land?
 - Selling fruits or vegetables?
 - Collecting firewood for sale?
 - Making charcoal for sale?
 - Making alcohol for sale?
 - Etc....
 - Are there any other activities you done in the past money to earn money?"
- "How much did you earn from (activity) during the past month?"
- "Did you receive money from any other sources that we haven't mentioned during the past month?"
- Have you received any money from relatives or friends in the past month? From whom? How much?

Consumption

By consumption, we mean the goods and services that are actually used during a given period.

Measures of consumption have several advantages over measures of income.

Consumption is more stable over time: households save and borrow to smooth their consumption over time; while we might see a large lump sum of cash income in one month then nothing for the rest of the year, we obviously would not see a large consumption of food in one month and then fasting for the rest of the year. This is not to say that consumption doesn't fluctuate over time though, and we'd still expect to see greater consumption during harvest time when incomes are higher or during festival periods. But, the fluctuations in consumption will be typically less than fluctuations in income and thus a better measure of well-being.

A second advantage of measuring consumption is that it is a more direct measure of well-being and can better capture how resources are divided within a household, for example with regard to differences between men and women or between adults and children.

An important distinction to make is between consumption, which refers to the goods and services people actually use, and expenditures—what they buy. Expenditures and consumption of perishable goods (e.g. meat or fruit) will be nearly identical—unless people produce their own. However, measures of consumption are better for more durable goods, since expenditures might occur in lump-sums in a similar fashion as income. For example, a family might buy food staples, such as rice or flour, in bulk. If you ask how much of this good they have purchased in the previous month, it doesn't tell you how often such purchases are made. In the long run, it should even out—if it is

random who buys what when—but it makes for a less precise understanding of well being.⁸

The downside is that consumption is even more difficult to measure than income or expenditures. For evaluations, monitoring or baseline surveys, it is unlikely that you'll have the resources to undertake a comprehensive questionnaire of all of the items that a household consumes. It's more realistic to make a list of specific items that are of interest or are most typical consumed by the population.

Food is the most common measure of well-being of poor individuals.⁹ While there is evidence that poor people by no means spend all of their income on food, it is a good starting place. We can ask how many meals a person or her children typically eat in a day to get a basic understanding of nutrition. We can go further by asking about specific types of food consumed and in what quantities. For example, "how many times per a week do you eat meat?" This gives an idea of quality of intake, which is more informative than quantity.

Asking about consumption of *luxury food items* is a good way of gaining an understanding of well-being. This will be context-specific and depend on the dietary habits of the local population. Some examples are sugar, cooking oil, and meat.

Alcohol and tobacco also give an idea of overall consumption patterns. While this information is useful, it should be noted that these items are often considered as sinful and the expenditures and consumption will most likely be *under-reported*.

Another important category of spending/consumption is entertainment. This, again, will be specific to the region in which the survey is conducted. Categories might include admission to entertainment events, for example projections of films or sporting events, purchases of lottery tickets or other money spent on gambling or airtime for mobile phones. Spending on festivals and other social events—including funerals and weddings—often constitute a large part of expenditures as well.

Consumption of *health* and *education* are very important in understanding the economic status of households, and interesting outcomes in their own right. We'll consider each of these in more detail below, however, asking about expenditures on health-care and education gives you an idea of how much income goes towards consumption in these categories.

⁸Deaton, Angus and Margaret Grosh, "Consumption Module," in "Designing Household Survey Questionnaires for Developing Countries, Vol 2," Grosh, Margaret and Paul Glewwe, ed., International Bank for Reconstruction and International Development, The World Bank, 2000.

⁹ Lower nutritional intake might be the result of lower well-being due to overall lower consumption, or it might be the result substituting food consumption for other items that are more valuable/enjoyable to consume. (Ibid).

- How many meals do you typically eat in a day?
- How many meals do your children typically eat in a day?
- How many meals do you eat meat in the past month?
- (How many kilos of rice (or locally appropriate food) do you eat in a typical week?
- How much salt/sugar... do you use in a typical week?
- How often do you use oil when cooking?
- How much have you spent in the past week on canned/packaged food?
- How much did you spend on motorcycle taxis (or other location-specific type of transportation) in the previous 7 days/month?

Assets

In Bartos and Lively (2014) “Livelihood development”, chapter 1 we discuss productive assets, which are used to generate wealth. Here are more concerned with ownership of items that increase well-being. Using assets as a measure of wealth has the advantage that they are easy to measure and people typically know what they own, which makes these measures more accurate than income, expenditure or consumption.

Assets change less with time. This is generally a good thing and makes it easier to measure. However, it also means that they will be less reactive to interventions, so even if programs are successful in increasing the income of participants, it might not be apparent in assets for a long time.

Usually assets are recorded at the household level since they are shared between household members. This means it's important to consider household size when making comparisons.

- What are the walls of your house made from? (e.g. wood, mud, stone, brick)
- What are floors of your house made from? (e.g. concrete, wood, linoleum)
- Does your household own the following items? (How many?)
 - Mobile phone
 - Radio
 - Television
 - Plates for eating
 - Bicycle
 - Mattresses
 - Chairs
 - Stove for cooking

Debt

As described in Bartos and Lively (2014) “Livelihood development”, chapter 2, there are various ways that poor households borrow and save money, even when they do not have access to formal banking. Saving and borrowing can both be used to smooth

consumption over time (as we discuss above), to finance expensive, "lumpy" purchases for consumption, or to buy productive assets that can be used to generate future income. In all three of these cases, getting a good idea of the debt held by an individual or household is essential to assessing their overall financial situation.

The same logic applies in asking about debt as with income: an itemized questionnaire of debt from a list of likely sources will help respondents to remember their outstanding debts.

It is also worthwhile to ask about the interest paid on outstanding loans. the easiest way to do this is to simply ask about how much money was borrowed from a given source, how much is expected to be returned and when payment is due.

Asking if scheduled payments have been late will also provide insight into the borrowing habits of the family and give you an idea if they have over-extended themselves.

Finally, informal loans between friends, family and neighbors are common in many settings. In addition to having borrowed money, it is possible that respondents will have lent money as well, and may be worth taking the time to collect this information.

- Are you (or anyone in your household) members of a ROSCA (use local term)?
 - If so, how often do you meet? How much do members contribute?
- Are you a member of a savings group or credit collective? How much savings do you have with this group?
- Do you owe any money to the following people? How much did you borrow? How much are expected to pay back? How long did you agree to borrow this money for? Have any of the scheduled payments been late?
 - Money lenders
 - Friends
 - Shop-keepers
 - Family
 - Micro-finance lenders
 - Village leaders
 - Anyone else

Health

Health is difficult to measure in simple survey questions. If you have enough resources, the best option is to collect direct measures of health, like weight and height. Measuring children's height and comparing it to their age is a particular useful measure of physical development. In addition, measuring the diameter of individuals' arms will give a rough indication of malnutrition.

When relying on survey questions alone, the goal is to ask about tangible behaviors or information that gives an approximation of health status. Dependent on the

environment, ask about specific diseases and health-related behavior. When appropriate, asking about sexual behavior (in a culturally sensitive manner, usually with better trained enumerators).

Asking about the number of visits to health clinics and spending on health care is informative, but doesn't necessarily capture actual health status, since people who are more conscious of health-related issues will visit clinics more frequently for the same ailments (i.e. if you observe that some people visit the doctor more often, it could mean that they have more health problems, or it could indicate that they have better access to health care or are better educated as to the benefits of health care).

- How many days have you missed from school/work in the past month due to health-related issues?
- Do you currently suffer from any health-related problems? What are they?
- Generally speaking, how would you describe your current health? (Very good? Average? Not very bad?)
- Would you say that you are healthier or less healthy than you were a year ago?
- Have you (your children) had diarrhea in the past month/year?
- Have you (your children) had a fever in the past month/year?
- How many bed nets does your household own?
- Do you/your children sleep under bed nets every night?
- Are you sexually active?
- How many sexual partners have you had in the past month/year?
- Do you use condoms regularly?
- Have you been tested for HIV? (If yes, are you HIV positive?)
- Are you covered by any health insurance? How much do you pay monthly? What is the deductible?
- How many times have you/household members visited health-care facilities in the past month/year?
- Have you/household members visited any traditional healers/mid-wives for health-care in the past month/year?
- Have you purchased any medicine in the past month? How much did you spend on medications for your household?

Infrastructure and sanitation

One way of approximating health is to measure not outcomes, but environmental factors or behaviors that are known to affect rates of disease and the physical development of children.

For example the type of water source used for domestic purposes can have an effect on rates of water-borne disease and thus health.¹⁰ Look at studies done in the country by

¹⁰ The World Health Organization (WHO) classifies water sources as "improved" or "unimproved" depending on characteristics of the water source. Tap water and boreholes or

the World Health Organization (WHO) or national demographic and health (DHS) surveys will give you some guidance on what typical water sources are and which are classified as improved.

Similarly sanitation facilities affect rates of disease. Classifications of toilet facilities are also provided by WHO and are included in DHS surveys. Toilets that are shared among multiple households also increase the spread of disease.

- What type of water source do you use for domestic purposes? (List options, for example: tap water, mechanical well/borehole, covered well, uncovered well, stream/pond).
- Do you take any steps to purify water at home? (List options, for example: boiling, filter through cloth, add chlorine/iodine).
- What type of toilet does your household use? (List options, for example: flush toilet, pit toilet covered with concrete slab, uncovered pit toilet, open defecation).
- Is this toilet shared with other households?

Education

Education is important to measure both in the context of human capital and socio-economic status, as well as in terms of spending and consumption of education. In the first category, asking about highest level of education of the respondent and their household members is a good starting place for human capital. It is also informative to ask about parents' highest level of education. The education of a person's mother is a particular good indicator of socio-economic status.

When education is of particularly low quality, however, years of education may be a poor measure of a person's true capabilities. For this reason, it is advisable to ask not only about years of education but directly about literacy with questions like:

- What is the highest level of education you have completed?
- What is the highest level of education completed by your husband/wife?
- What is the highest level of education completed by your father (mother)?
- Are you able to read a book? If no, are you able to read a poster or notice?
- Are you able to write a letter?
- Which of your children are in school?¹¹
- (If a child does not attend school) Why is this child not in school?
- How much did you spend in the previous year on school fees and uniforms?

When asking about school enrollment of children it is often instructive to enquire about why children do not attend school, for example, whether it is because they are needed at

mechanical wells are considered improved, and correlated with less disease. Uncovered wells or open water are correlated with much higher rates of disease.

¹¹ You might ask about number of children and their ages, then have enumerators fill out a table with school enrollment for each child, along with their age and reason for not attending school.

home, because school fees are prohibitively expensive or because of attitudes towards schooling for girls.

Choosing survey questions

In this chapter, we have presented some possible ways of eliciting information on income, consumption, assets, health, and education. There are, of course, many more questions and detailed information that might be collected about each of these categories. Which questions are included on the survey will depend on what information is most relevant to the study you are conducting.

Key reading

Deaton, Grosh (2000). Designing Household Survey Questionnaires for Developing Countries: Lessons from 15 years of the Living Standards Measurement Survey. The International Bank for Reconstruction and Development, The World Bank, Washington DC.

© People in Need, 2014

Data collection: A practical guide to collecting data in developing countries

Vojtěch Bartoš and Ian Levely