

Authors
Alex Bartik

Why getting good data is really hard...

We just completed the baseline survey for one of the projects I'm working on in Cagayan de Oro, the Philippines. Consequently, I've been spending most of my time the last few weeks thinking about how to get good data. Getting the right information might seem simple; figure out what you want to know about people and then ask them. However, in practice getting good data proves much more difficult.

First, even if you know what you want to know, figuring out how to actually ask the right questions remains very difficult. We're usually doing surveys at people's houses and we have limited time (both because of surveyor costs and because people won't complete an excessively long survey) and consequently we often can only devote a few minutes to each topic. For example, for the Informative Advertising and Spillovers project we wanted a basic measure of financial literacy. However, we really only had space for 3 or 4 questions on financial literacy. Clearly, we cannot measure all aspects of financial literacy well with only 3 questions. Therefore we had to carefully figure out what exact financial knowledge or skills we wanted to test and calibrate our questions to ask exactly those skills or items.

Second, getting people to agree to complete a survey is not always easy; people are busy and sometimes distrust the motives of people asking them questions. In Cagayan de Oro, the residents of one neighborhood were extremely hesitant to answer our questions because they had a bad experience with an NGO that had previously surveyed the area. That NGO had not gotten the requisite approval from the local authorities and people had not been happy with what they had done (I'm still not clear on the exact reasons for this unhappiness). We had the appropriate permissions from the authorities, but I had not provided each surveyor with a signed copy of the approval, and consequently the surveyors had no hard proof that we were there legitimately.

Finally, survey conditions can make obtaining accurate answers difficult. For the Informative Advertising and Spillovers project we have some math questions designed as a very rough measure of cognitive ability. To my surprise, many people enjoy these questions. However, they enjoy them so much that often times people's neighbors or family members will crowd around us while doing the survey and try to help with the math questions. We appreciate the enthusiasm, but we want to know whether the person we're talking to can answer the questions, not if they have a neighbor who can do so.

Making sure we get good data is not the most glamorous part of being a Project Associate; it involves relentlessly making sure that every question is clear both to you, the surveyors and the respondents as well as making sure that every survey is properly recorded and stored.

However, I really think its one of the most important parts of my job. In order to figure out what works and what doesn't we need some good data first.

June 19, 2009