

Measuring Aspirations: Discussion and Example from Ethiopia

Tanguy Bernard and Alemayehu Seyoum Taffesse***

** Research Fellow, Markets, Trade, and Institutions Division, International Food Policy Research Institute*

*** Senior Research Fellow, Development Strategy and Governance Division, International Food Policy Research Institute – Ethiopia Strategy Support Program II, Ethiopia*

This paper is also published as IFPRI Discussion Paper 01190; Markets, Trade and Institutions Division; Development Strategy and Governance Division

TABLE OF CONTENTS

Abstract	1
1. Motivation and Purpose	1
2. Measuring Aspirations.....	2
2.1. What Is Measurement?.....	2
2.2. What Are Aspirations?	2
2.3. How to Measure Aspirations	3
3. Toward an Instrument to Measure Aspirations.....	4
3.1. Related Literature	4
3.2. Definition of an Instrument	6
4. Testing for Usability, Reliability, and Validity of the Aspiration Measure	7
4.1. Usability	9
4.2. Reliability	9
4.3. Validity	12
5. Conclusions	14
References.....	15

LIST OF TABLES

Table 4.1—Reliability 1: Test-retest and observer reliability.....	10
Table 4.2—Reliability 2: Mood dependence	11
Table 4.3—Reliability 3: Anchoring effects.....	11
Table 4.4—Reliability 4: Wording	11
Table 4.5—Weights	12
Table 4.6—Determinants of aspirations	13
Table 4.7—Aspiration-related behavior	14

LIST OF FIGURES

Figure 4.1—Layout of aspiration questionnaires	8
--	---

ABSTRACT

Individuals' aspirations and their consequences for future-oriented behavior have received increased attention in development economics literature in recent years. At this stage, however, each study relies on ad hoc empirical instruments to measure aspirations, thereby limiting comparability of the results obtained. This paper proposes a simple measurement instrument that spans several dimensions aggregated via individual-specific weights. We use a purposefully collected data set to test for the usability, reliability, and validity of the instrument. In addition to standard test-retest procedures, our innovation lies in the use of several randomized tests introduced within the questionnaires themselves, in the enumerators' qualifications, and in the information set available to respondents. Overall results show strong support for the proposed instrument, with the caveat that collection of such attitudinal data requires experienced enumerators capable of adequately probing respondents.

Keywords: aspiration, measure, reliability, validity, Ethiopia, economics literature, data, poverty, survey methods, sampling methods, microeconomic analysis, microeconomic behavior

1. MOTIVATION AND PURPOSE

Following recent work by Arjun Appadurai (2006) and Debraj Ray (2006), several papers have attempted to empirically investigate the intertwined relationship between aspirations and poverty (for example, Macours and Vakis 2009; Bernard, Dercon, and Taffesse 2011; Beaman et al. 2012; Knight and Gunatilaka 2012). The indicators used to characterize aspirations vary substantially from one study to another, including depression scales, positive feelings toward the future, locus of control, goals, and more. All may be positively related to one another and to the idea of aspiration in general and thus may serve as adequate proxies to relate aspiration and future-oriented behavior. However, when it comes to the study of aspirations per se and the processes that lead to their formation, one may need to distinguish between the different types of mental models and states to which they relate. Further, to facilitate comparison of results across upcoming studies,¹ one may wish to rely on a standardized empirical definition of aspirations and related concepts—as an analogy, think of the considerable progress that has been made in poverty comparisons since the Foster-Greer-Thorbecke (1984) indicators.² The objective of this paper is to provide a first step toward such an indicator by proposing a simple measurement instrument that can eventually suit this purpose.

Defining an empirical instrument for aspirations is not without its challenges. Although long used in all other social sciences, attitudinal indicators have traditionally raised skepticism among economists, primarily due to issues of anchoring, wording, and scale dependence; respondent role-playing; and instability of respondents' moods over time. These may in turn affect interperson comparability of responses, as respondents may understand questions differently, or even intrapersonal comparability, as respondents may interpret the same wording differently when directed toward different aspects of their lives (see, for instance, Bertrand and Mullainathan 2001 or Manski 2004 for reviews and discussions). Such measurement errors may raise important issues when used for econometric inference, the most severe occurring when the attitudinal indicator considered is used as the dependent variable in estimation (Bertrand and Mullainathan 2001)—for instance, when one wants to study the formation of aspirations. Overall, while increasingly recognizing the need for better understanding of choice processes, the profession remains more inclined to believe what people do, not what they say.

However, as argued in Manski (2004), in the absence of information regarding why and how people make their choices, one is left to assume a particular type of preferences, expectations, or both when analyzing choice data. Yet the combination of different types of preferences and beliefs could lead to the same observed choices but imply different policy consequences. Thus, several recent studies have attempted to assess the reliability of such attitudinal data as expectations or satisfaction with life measures and their possible limitations (for example, Manski 2004; Delavalande, Giné, and McKenzie 2011; Krueger and Schkade 2008). The broad answer is that provided enough care is taken at the time of designing the instrument, relevant analysis can be performed to usefully inform researchers about individuals' decision-taking processes.

This paper proposes a similar approach to the design of a measurement instrument for aspirations. Specifically, we review the definition of *what aspirations* are and what they're not and derive a simple empirical definition of *aspirations*. Using two rounds of purposefully collected data in rural Ethiopia, we then test for the usability, reliability, and validity of the indicator defined. Specifically, reliability is assessed through test-retest, variation in enumerators' experience, change in respondents' moods, random variations in the ordering of questionnaire modules, and relationship with scale-based

¹ And to avoid the sort of data mining whereby a wide variety of indicators are collected but only the statistically significant ones are reported.

² See Foster, Greer, and Thorbecke (2008) for an assessment of contemporaneous use of their indicators.

measures. Validity is assessed by providing extra information to randomly chosen respondents and by testing for construct validity in both the determinants of and the future-oriented behavior predicted by different levels of aspirations.

Overall, our results provide strong support for the reliability of the instrument's different components. A significant caveat highlights the need for experienced enumerators capable of adequately probing respondents. Further, the designed instrument performs well—arguably better than others—in its capacity to predict future-oriented behavior, as well as with the expected determinants of aspirations. The instrument will be used later in an experimental study investigating the process of aspiration formation. To our knowledge, the present paper provides an original attempt in economics at rigorously testing for the validity and reliability of an attitudinal measure before it is used as a main indicator in an actual experiment.

The rest of the paper is organized as follows. Section 2 briefly reviews issues of measurement and their application to the concept of aspirations. Section 3 reviews recent studies that provide empirical assessments of determinants and relevance of aspirations in explaining future-oriented behavior based on a set of diverse measurement tools. We then propose a measurement tool to be tested for its usability, reliability, and validity in Section 4. Section 5 concludes.

2. MEASURING ASPIRATIONS

2.1. What Is Measurement?

Before turning to the actual design of a measurement instrument for aspirations, it is useful to briefly recall the properties expected from a measurement tool. Measurement is a systematic, replicable process by which objects or events are quantified, classified, or both on a particular dimension. This is usually achieved by the assignment of numerical values (Weiner 2007). Although simple in appearance, this constitutes a rather complex and controversial process; indeed, the why and how associated with measurement have led to the development of different measurement theories. Although it is not the intention of this paper to explore the rather large literature on measurement in statistics, psychology, and other disciplines,³ several issues are worth highlighting for our purpose.

The first question relates to the level of measurement. Four such levels or types are identified—nominal, ordinal, interval, and ratio. The second question focuses on the aim of measurement. The two main strands noted are representational measurement and operational measurement. Hand (1996) describes them as follows. “One way of thinking about the difference between representational and operational measurement is that the former seeks to *represent* or *model* empirical relationships—and so is about *understanding* the substantive domain of investigation—whereas the latter seeks to *predict*. Accurate prediction can be achieved without any understanding of the underlying mechanism (witness someone who can drive a car well without any understanding of how it works). Confusion between these two aims is widespread.”

A third relevant dimension is concerned with what may be characterized as properties of measures. Two such attributes are highlighted—reliability and validity. Reliability of a measure refers to the degree to which a measurement technique can be depended on to secure consistent results on repeated application, whereas the validity of a measure is the degree to which any measurement approach or instrument succeeds in describing or quantifying what it is designed to measure.⁴ An additional attribute may be added relating to an indicator's effective usability, that is, how difficult it is to implement and whether it can be done in standard surveys.

In short, it is desirable to identify and/or develop measures of aspirations and related concepts that are consistent with the purpose of the analysis (descriptive or analytical), reliable, valid, and usable.

2.2. What Are Aspirations?

As previously noted, measurement is about assigning numerical values to objects. The process thus begins with the definition or characterization of the object of interest, which in the present case are aspirations.

A dictionary definition of the word *aspiration* is “a desire or ambition to achieve something” (*Oxford English Dictionary* 1989). The word thus signifies some aim or target and a desire or wish to attain that goal. The meaning also suggests, rather implicitly, that some effort would be exerted to realize the desired aim/target. Moreover, an aspiration may or may not reflect the feasibility of the corresponding target. In short, aspirations combine or summarize the preferences main-

³ For further details see Hand (1996) and Blanton and Jaccard (2006).

⁴ Exploring the nature of reliability and validity as well as developing ways of testing whether measures have these properties led to a large body of work. See, among others, DeVon et al. (2007) and Weiner (2007).

tained, the beliefs held, and possibly the constraints acknowledged by an individual about aspects of the future. The key implication is that aspirations can influence an individual's future-regarding behavior.

The following descriptions aim to further clarify the concept. Preferences are the attributes of individuals, other than beliefs and capacities, that constitute the reasons for behavior, that is, account for the actions they take in a given situation (Bowles 2004).⁵ Beliefs are about the structure and details of the world we experience (and potentially other worlds) and the implied relationship between actions and outcomes and about the behavior of others in various contingencies, whether or not the behavior actually occurs. Thus, beliefs reflect the mental models that individuals develop and/or learn to explain and understand their environments.

In short, aspirations summarize a subset of an individual's beliefs, preferences, and capacities that are specifically relevant to behavior regarding the future.⁶ The latter three are in turn symbiotically related to the institutional configurations within which the individual functions and are conditioned by exogenous shocks. This way of looking at aspirations affords a number of advantages. First, it clarifies the nature of aspirations vis-à-vis beliefs and preferences. Aspirations are not preferences or beliefs per se. They represent those beliefs and preferences that govern the future-regarding behavior of an individual. Second, it highlights the role of aspirations as determinants of—or as motives and reasons for—behavior, including dysfunctional aspects of behavior. Third, it is consistent with the conventional approach (practice) of explaining behavior on the basis of beliefs, preferences, and capacities. As such, it can help organize the various explanations offered and distinguish the focus of the present research (see Bernard, Dercon, and Taffesse 2011 for a discussion). Finally, it provides a convenient way of introducing cultural determination of aspirations since all beliefs, preferences, and capacities are conditioned, but not necessarily fully and exhaustively determined, by social circumstances. These include, among others, culture, history, level of development, and political structure.

2.3. How to Measure Aspirations

Three observations need to be made at this point, each referring to a particular feature of aspirations. First, a potentially large set of determinants may be associated with aspirations. For instance, following the words of Ray (2006), aspirations are rooted in one's aspiration window—that is, one's cognitive world, one's zone of similar, attainable individuals on whom one relies to assess what is feasible for oneself—which is influenced by the size and composition of one's network of contacts. Yet aspirations are also conditioned by the personal life experience of the individual.

Second, aspirations, like all attitudinal traits, are not directly observable. It is not possible to observe aspirations in the same way that one can observe wealth. Two possible alternatives may be devised. The first is indirect, combining assumptions about the set of aspiration determinants or the pattern of choices made by individuals and their attendant outcomes with data on actual choices. For example, a person with a narrow aspiration window would be expected to have a low level of aspirations. Thus, low investment on the part of isolated individuals provides an initial indicator of limited aspirations. However, such an approach clearly relies on strong assumptions about the role played by aspirations and their determinants, and it is likely that alternative explanations for limited investments could be found without adequate means to test them against an aspiration perspective (see Manski 2004 for a similar discussion about preferences, expectations, and choice outcomes).

The second alternative is to ask individuals directly about their aspirations. This approach can generate good-quality data if implemented carefully. However, two sources of bias in self-reported measures should be noted. One set relates to participants and includes participants' willingness to report private knowledge, evaluation apprehension, and participant role-playing. The other concerns the various attributes of the instrument used and includes prior questions (anchoring), the number of categories on the rating scales, the adjectives that are used, the scale endpoints, and the adverbs that describe scale categories. In the best-case scenario, such measurement errors are of a classical type and lead to only attenuation biases in estimations, whereby aspirations would be used as an independent variable, with no bias if used as a dependent one. Yet when correlated with the level of reported aspirations per se, measurement errors will lead to biased results in both types of estimations (Bertrand and Mullainathan 2001). As with any attitudinal data, an adequate direct measure of aspiration must therefore satisfy a number of preliminary tests before any further analysis can be performed—a well-known feature in psychological literature where tests for reliability and validity have long been implemented on similar attitudinal concepts.

Third, like well-being itself, aspirations span multiple and potentially interrelated dimensions. Individuals may have wealth or income aspirations, educational aspirations, social status aspirations, or aspirations for others such as their children. Although in principle it is possible to use constituents, determinants, or both to assess the level of aspirations, this can be

⁵ Note that norms form a part of preferences. Greif (2005) develops a similar characterization of beliefs in the context of institutional economics.

⁶ It is possible to view the "development as self-discovery" characterization (Hausmann and Rodrik 2003) at the individual/community level from this perspective.

done for only a single dimension. Nevertheless, in some cases, the measurement of aspiration levels may be more effectively achieved via the construction of a summary measure such as an aspiration index.

3. TOWARD AN INSTRUMENT TO MEASURE ASPIRATIONS

Several recent papers provide empirical insights into the determinants and the role played by aspirations in shaping future-oriented behavior. We briefly review their results and the instruments used to characterize aspirations before proposing a new instrument in Section 4.

3.1. Related Literature

Macours and Vakis (2009) investigate how social interactions affect aspirations in drought-prone poor communities of northwestern Nicaragua. They rely on the random grouping of beneficiaries into three variants of a conditional cash transfer program. By design, each group encompassed a random variation in the number of individuals with leading positions in their communities. Because the program induced regular meetings among group members, one of the effects was to foster social interactions among beneficiaries within a group. Their results show that the greater the proportion of natural leaders within one's group, the more likely one was to report a positive attitude toward the future—particularly for those groups benefiting from the largest transfers in the program. Further, individuals within these groups were more likely to report higher income and higher investments in children's schooling and nutrition. Overall, their results suggest that such positive attitudes indeed may be contagious in that witnessing local success stories of upward mobility can be important in changing individuals' investment behavior.

Macours and Vakis (2009) rely on a widely used measure of depression referred to as the Center for Epidemiological Studies Depression Scale (CESD). The CESD asks respondents about the occurrence of 20 mood traits during the preceding week and aggregates the results following a predefined scale. Its purpose is to provide shorthand evaluation of the severity of depression that individuals may be going through. In their paper, Macours and Vakis used both the aggregate CESD indicator, along with the sum of scores on positive and negative feelings-related questions, and the separate answer to questions of being cheerful and having strong positive expectations toward the future. Finally, they asked respondents whether they would or would not do anything to prevent the impact of future shocks. All indicators perform relatively well in their analysis, converging toward the idea that social interactions indeed did enhance individuals' positive self-perception. Yet one may argue that although related to aspirations (either as codetermined by other factors or as the result of one another), depression scales or positive feelings toward the future may be qualitatively different from aspirations per se. As an example, Ray (2006) argues that an aspiration gap—the difference between one's actual level and the level he or she wishes to attain—is related with effort through an inverted U shape, for a low aspiration gap may relate to resignation and a high aspiration gap to frustration. Both states could fit with a sense of depression.

In another study, Bernard, Dercon, and Taffesse (2011) attempt to justify the use of an aspiration failure framework to describe the type of "fatalistic" behavior that often characterizes poor Ethiopians. Their analysis relies on cross-sectional data, and their results are limited to assessing whether correlations predicted by the aspiration failure framework actually exist, namely, that a narrower aspiration window should lead to lower aspirations and that a lower aspiration gap should lead to underinvestments. Their results confirm the relevance of such concepts in understanding the role of fatalism and its implications for future-oriented activities such as borrowing for new business activities or children's education. Understood in this sense, the evidence suggests that fatalism lowers the demand for long-term loans and the use of these loans for productive purposes.

The main indicator used in Bernard, Dercon, and Taffesse (2011) rests on two dummy indicators akin to those used in the literature interested in locus of control concepts. Specifically, respondents were asked to answer the following questions:

For each of the following, please tell me which of the two propositions you most agree with

- A.
 1. "Each person is primarily responsible for his/her success or failure in life."
 2. "One's success or failure in life is a matter of his/her destiny."
- B.
 1. "To be successful, above all one needs to work very hard."
 2. "To be successful, above all one needs to be lucky."

Here again, although the study relies on a well-tested and -documented indicator that seems positively related to aspirations (those who report an external locus of control being more likely to have a lower aspiration level), the authors cannot ascertain that it is aspiration per se that is being captured independently of other relevant psychological attributes—for instance, self-efficacy. In fact, their discussion suggests that aspirations, locus of control, and self-efficacy,

although similarly related to “relevant others” and similarly predicting future-oriented behaviors, remain qualitatively different. Further, as argued in Manski (2004), the use of a binary indicator to capture differences in a continuous latent attitude is prone to important margins of error.

In a recent study, Beaman, Dulfo, Pande, and Topalova (2012) investigate the effect of the West Bengal reservation rule for Pradhan⁷ positions on boys’ and girls’ aspirations. At each election since 1998, one-third of these seats, chosen randomly, have been reserved for women. The authors find that compared to nonreserved councils, in councils in which the Pradhan was selected to be a woman, parents have significantly modified—upward—their aspirations for their girls; no such increase was found for boys. These results are confirmed by similar questions asked of adolescent girls and boys within these families.

To measure aspirations, Beaman et al. (2012) took a rather direct approach and asked respondents (1) the level of education they wished (their children) to attain, (2) the type of occupation they wished (their children) to have at age 25, (3) the age at which they wished (their children) to marry, and (4) whether they wished (their children) to occupy a Pradhan position in the future. Based on the answers, they computed a set of five dummy indicators: graduation from high school or above; any occupation different from housewife or what is required by the in-laws; career as a doctor, scientist, teacher, or lawyer; married at age 18 or older; and desire to become a Pradhan. They find strong positive correlations between the first four dummies but not the last one; they then use this result to exclude the last dummy indicator in a standardized average index of aspirations that spans the first four. In fact, because these indicators represent different spheres of aspirations that may be differentially affected by the presence of women in leadership positions, aggregating the five variables altogether “may not be legitimate and may confound the effect of reservation onto the various spheres of aspirations.” The Beaman et al. (2012) paper thus offers a rather straightforward attempt to measure aspirations. In fact, the questions the authors rely on are akin to those used in most literature dealing with aspiration and educational attainment. Yet as for any of the attitudinal indicator, the reliability and validity of the individual components and their aggregate, still remains to be tested. Further, because the wording of the questions differed across indicators (for instance, aspired education level was assessed with the number of years of schooling achieved, whereas the occupational question was an open-ended question that led to categorical sorting), their translation into dummy indicators was needed for aggregation purposes, considerably reducing the exploitable variance in the data collected. Finally, although acknowledging that aspirations span different dimensions, in the absence of a respondent-specific weighting scheme, their aggregate index could not include part of the information collected.

Last, Knight and Gunatilaka (2012) rely on two questions developed by Stutzer (2004) in Switzerland to assess variations in income aspirations in China. Results indicate that aspirations evolve positively with previous income, as well as with peers’ income, but are negatively related to one’s self-assessment of well-being. Specifically, their measure of aspirations relies on the following two questions:

1. “What income would you indicate as good or bad in your circumstances? Please try to state what income per month (before taxes) for your entire household you consider to be ... sufficient.”
2. “What household income per month would you consider an absolute minimum in order to make ends meet and without running into debt? We do not only mean housekeeping allowance but all essentials, including insurance, taxes and so on.”

Important to note, Stutzer (2004) remarks that current households’ income is in general higher than that reported in question 2 and also often is higher than that reported in question 1. Accordingly, “This might indicate that the proxy measures do not capture people’s aspirations.” He notes, however, that the indicator only needs to be sufficiently correlated with the “true” aspiration level to test the propositions set forth.

Several conclusions can be extracted from the four studies described here. First, albeit in different contexts (Nicaragua, Ethiopia, India, and China), all studies converge toward the importance of aspirations in influencing future-oriented behavior—whether it is school enrollment, nutrition, or other future-enhancing investments. They also support the idea of Appadurai’s (2006) aspiration windows, whereby aspirations are positively influenced by “relevant peers” in one’s community. Altogether, these features provide substantive support to the idea that the relationships described exist over and above issues of measurement errors highlighted in the economic literature, thereby calling for further empirical investigation into the role played by aspirations in affecting well-being as well as the process by which aspirations are formed.

Second, each study relies on different measurement tools. Some rely on well-tested indicators that enable generic formulation to capture people’s general mindset. However, these indicators were not specifically designed to measure aspirations per se and, although closely related, may capture psychological traits that differ from aspirations in nontrivial

⁷ Leader of village-level council.

ways. Others rely on a straightforward assessment of individuals' aspirations, with the limitations that the reliability of such questions remains to be assessed and that the multidimensional features of aspirations are only weakly taken into account. Thus, further research on the process of aspiration formation and aspirations' consequences for individuals' behavior calls for the elaboration of a new instrument.

3.2. Definition of an Instrument

Based on the foregoing discussion, we propose a simple instrument to measure aspirations. In general terms, our instrument is akin to that of Beaman et al. (2012), relying on direct questioning of respondents' aspirations for various dimensions of their lives, with the added feature that the same simple wording is used for each dimension to facilitate later aggregation. We describe these wordings and corresponding scales below.

Important to note, when asked such types of questions directly, individuals are likely to report their general wishes instead of the actual subset of an individual's beliefs, preferences, and capacities that are specifically relevant to behavior regarding the future, per the definition of aspirations in the section entitled, "What Are Aspirations?". In essence, this is akin to the type of overconfidence in probabilistic reasoning reported by Morgan and Henrion (1990). Focusing on expectation, Manski (2004) thus suggests that individuals first be asked for their assessment of the minimum and maximum level possible for each dimension before being asked about their actual point beliefs. This has the added value of reducing potential anchoring effects related to previous questions. Therefore, respondents shall first be asked about possible minimums and maximums in their community for each dimension of life in which we seek to measure aspirations.

Although respondents are allowed to provide answers then for their own aspired level above the maximum reported, these questions are deemed relevant to help them frame their answers. To further help respondents answer such unconventional questions, visual scales are used, featuring high (but reasonable), medium, and maximum levels obtainable for each dimension in the survey area. Here again, the purpose is merely one of framing the questions; respondents shall be free to report levels higher than those proposed. For dimensions such as income or wealth, these medium and maximum levels are also translated into meaningful examples, helping respondents to translate their aspiration levels into financial values.

Finally, to further facilitate respondents' answers and because it also may provide a useful quantity for analysis, respondents are asked about their present level. Overall, we propose that the set of questions used for each dimension chosen to characterize individuals' aspirations be asked in the corresponding order:

- A. *With respect to (dimension k):
What is the maximum level of (dimension k) that one can have in your community?
What is the minimum level of (dimension k) that one can have in your community?*
- B. *With respect to (dimension k):
What is the level of (dimension k) that you have at present?*
- C. *With respect to (dimension k):
What is the level of (dimension k) that you would like to achieve in your life?*

To produce an aggregate index of aspirations, responses to question C must be comparable across the various dimensions. This is facilitated by the use of the exact same wording for each dimension; however, the scales and distribution still may differ across dimensions. Thus, the first step consists of standardizing respondents' answers by removing the sample mean from each observation and dividing it by the standard deviation of the said dimension in the sampled population. Each dimension is then dimension free and can be used readily in an aggregated index with other dimensions.

A further issue is the importance that one allocates to a particular dimension of his or her life outcome. With heterogeneous preferences, some respondents may value social status within their community more than their level of wealth, whereas it may be the opposite for others. Although all respondents may report high aspiration levels for both dimensions, unless forced to reveal their idiosyncratic preferences, the aggregate indicator will not capture these distinctions—a feature that is well-known in marketing research and that led to the development of conjoint analysis methodologies.

Therefore, we propose that after providing answers to questions related to each dimension, respondents are asked to weight each dimension according to the value they attach to it. This can be easily done by a question such as

- D. *Now I would like you to tell me which of these K dimensions are the most important for you. Here are 20 coins. Please distribute all 20 coins across the K aspects that we have discussed, according to their importance. No coin in a category means that you do not attach any importance to it. Many coins in a category mean that you attach a high importance to it.*

Let a_i^k be individual i 's response to question C above applied to dimension k , and let w_i^k be the weight that individual i assigned to this dimension when answering question D. The aspiration index can thus be expressed quite simply as

$$A_i = \sum_k \left(\frac{a_i^k - \mu^k}{\sigma^k} \right) \cdot w_i^k,$$

where μ^k and σ^k measure the sample mean and the standard deviation for responses to question C, respectively.

4. TESTING FOR USABILITY, RELIABILITY, AND VALIDITY OF THE ASPIRATION MEASURE

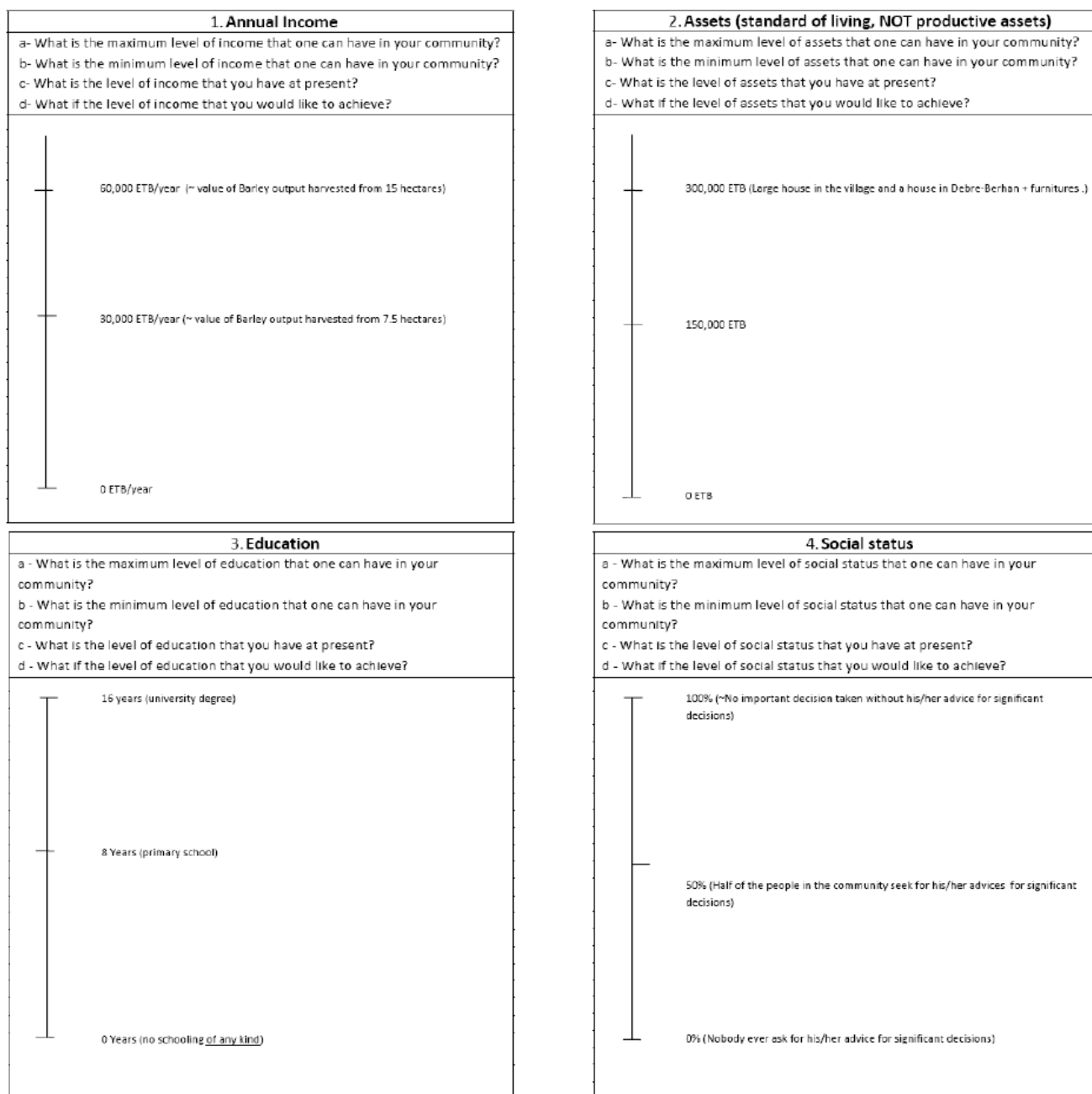
This section outlines the empirical strategy used to test for the adequacy of the proposed instrument to measure aspirations. Specifically, the instrument was assessed along the following lines:

1. Usability—Can the instrument be administered within standard surveys? Are respondents willing to answer such questions?
2. Reliability—Can the instrument be trusted to provide consistent measures of aspiration on repeated applications? To what extent are the obtained answers conditioned by enumerators' capacities, the questionnaire design, or both?
3. Validity—Is the instrument effectively measuring only aspirations? Are the obtained responses in line with expected determinants of aspirations and corresponding future-oriented behavior?

To this end, a purposely designed pilot survey was implemented in 16 villages of central Ethiopia, clustered around the town of Debre Berhan. The sampled individuals were all members of households repeatedly interviewed within the Ethiopian Rural Households Survey (ERHS) during the past 20 years. The use of ERHS data enabled purposeful sampling of individuals spanning different genders, education levels, and ages. Further, it enabled us to form pairs or triples of individuals living within the same community and sharing similar age, gender, and educational characteristics. For this purpose, we constructed 60 sample categories of individuals, each one being a unique combination of various levels of education (five levels), gender (two levels), and age (six levels). Within each village, we then selected individuals for whom we could find at least one other individual belonging to the same category within the ERHS sample. For each selected individual, our sample therefore includes at least one peer living in the same village, belonging to the same age group, with a similar level of education, and of the same sex. Overall, the survey covered 170 individuals within 16 villages.

The questionnaire was short and designed for the sole purpose of this pilot assessment of aspiration measures. It included aspiration-related questions as described above spanning four dimensions: yearly income, wealth (nonproductive assets), education, and social status within the community—the latter being captured by the proportion of individuals in the village who would ask for one's advice before taking an important decision. Figure 4.1 presents the layout of the aspiration-related module used. Important to note, these questions were asked for the respondent him- or herself as well as for his or her peer, defined as that other person within the same village who shares similar age, gender, and educational characteristics.

Figure 4.1—Layout of aspiration questionnaires



5. Now I would like you to tell me which of these four dimensions are the most important for you. Please distribute all the twenty coins in the four squares above according to their importance. No coin in a category means that you do not attach any importance to it. Many coins in a category means that you attach a significant importance to it.

a. Number of coins for "income"		b. Number of coins for "assets"		c. Number of coins for "education"		d. Number of coins for "social status"	
---------------------------------	--	---------------------------------	--	------------------------------------	--	--	--

Source: Authors' creation.
 Note: ETB = Ethiopian birr.

The questionnaire further included questions related to the respondent's present mood and locus of control perception. It also featured Likert-type scales related to the various aspiration dimensions, questions about respondents' exposure to information outside the community, and hypothetical questions related to one's demand for loans.

To assess the reliability of responses and to further introduce tests for validity, two rounds of the survey were administered two weeks apart using the same questionnaires. The choice of two weeks was motivated by the need for relative stability of the aspiration-related levels and at the same time ensured that respondents did not clearly remember the responses that they had given in previous rounds.⁸

⁸ Other studies of reliability often rely on the time span between the two surveys ranging from a week to two months.

4.1. Usability

The first question one asks when designing such an instrument concerns its actual usability. To what extent, for instance, are individuals willing to answer such questions? How difficult is it for enumerators to administer the questionnaire? Although enumerators did report the need for in-depth explanations of the core aspiration-related questions (see Figure 4.1), the use of visual scales and the anchoring within known levels of possibilities seem to have helped with this matter. Overall, 10 to 15 minutes were necessary to answer these questions, allowing for the use of such questions within broader surveys. At a minimum, the response rate obtained for each aspiration-related question (including weights) is greater than 95 percent.

We further asked the survey supervisor to provide us with a succinct report on the survey implementation. His main conclusions hint at the validity of the instrument in that it already provides elements of validity: “Generally speaking, the respondents participated well and the whole field work of the survey was successful. Most of them do aspire about their future life. They want to be richer as well as having more education. The aspiration was mostly age dependent. The inclination for aspiration for old age was not this much due to weak physical conditions and hopelessness. The middle age people aspire more about income and asset accumulation and less to education while the youngsters aspired first to education and then to income and asset accumulation.” Overall, from response rates and the supervisor’s reports, the usability of the designed instrument does not seem to pose any major issues.

4.2. Reliability

Because aggregation may lead to error reduction, all tests for reliability proposed here are performed on the individual components of the proposed aspiration index—that is, answers to questions C and D described in the Definition of an Instrument section for each of the four dimensions considered: income, wealth, education, and social status.

As a first step, we assess reliability via the most common approach found in the literature, which compares the correlation coefficient between an individual’s response across two rounds of surveys—the so-called test-retest procedure. To provide a more thorough understanding, we further investigate for reliability in relation to other dimensions. The first is referred to as observer reliability; namely, to what extent is reliability dependent on enumerators’ qualifications? Anyone conducting field surveys is aware of the importance of good-quality enumerators. With such attitudinal questions as aspirations, the stakes may be even higher. Thus, we implemented both surveys using four enumerators of two types, each being allocated a sample of 40 to 50 households depending on the village-level intensity of sampling. The first type consisted of experienced enumerators who had actually been involved in some of the previous rounds of ERHS data collection within the communities where the current survey was implemented. The second type was inexperienced enumerators. Further, for the second survey, the sample surveyed by one of the experienced enumerators in the first round was attributed to an inexperienced enumerator, whose sample was then attributed to the former. Together, test-retest and change of enumerator for part of the sample enable us to assess reliability across rounds, enumerator type, and change of enumerator.

We report the results of these tests in Table 4.1. Column 1 reports the reliability ratios between round 1 and round 2 surveys in the entire sample. The results indicate clear positive relationships between the two rounds of surveys, although with sometimes limited levels of correlation. As a benchmark, the typical reliability ratios in earnings and education measures used in the literature range from 70–92 percent and 80–94 percent correlation, respectively (Angrist and Krueger 1999). In some of the most commonly used attitudinal data, such as satisfaction with life indicators, a recent study finds a magnitude of around 60 percent (Krueger and Schkade 2008). Thus, although our measures of wealth and education perform relatively well, income, social status, and weight associated with the various dimensions can be considered weak.

Table 4.1—Reliability 1: Test-retest and observer reliability

	Whole Sample (1)	No Change of Enumerator between Round 1 and Round 2			With Change of Enumerator (Enumerators 3 and 4)
		Enumerator 1 (2)	Enumerator 2 (3)	Total (4)	Total (5)
<i>Income</i>					
Actual level	55.89*	35.57*	98.08*	55.07*	60.49*
Aspired level	33.14*	16.02	93.94*	45.51*	4.38
<i>Wealth</i>					
Actual level	82.29*	72.19*	99.10*	89.73*	79.52*
Aspired level	62.81*	59.15*	77.04*	69.72*	51.37*
<i>Education</i>					
Actual level	96.42*	99.26*	1.00*	99.71*	90.62*
Aspired level	92.10*	80.26*	97.92*	90.92*	94.05*
<i>Social status</i>					
Actual level	59.92*	36.48*	96.25*	82.81*	33.72*
Aspired level	20.87*	17.55	79.23*	65.74*	-9.98
<i>Weights</i>					
Income	40.26*	19.44	74.96*	48.49*	16.84
Assets	19.15*	31.07	72.74*	51.54*	5.27
Education	51.25*	56.70*	84.05*	74.23*	15.01
Status	49.07*	-1.35	56.08*	32.90*	13.38

Source: Authors' calculations.

Note: Reported are correlation coefficients between round 1 and round 2 answers. *Significant relationship at 5 percent level or less.

In columns 2 and 3, we report the reliability ratios for enumerator 1 (inexperienced) and enumerator 2 (highly experienced), respectively. The results are rather clear, showing much higher reliability ratios for enumerator 2, with orders of magnitude well attuned to those of non-attitudinal data. In contrast, the less-experienced enumerator performs rather poorly. Column 5 further exploits the variation introduced between rounds 1 and 2, where enumerator 3 (inexperienced) and enumerator 4 (experienced) switched samples. Although we cannot report reliability ratios for these enumerators separately, the comparison of results in column 5 with those in columns 2 and 4 provides some insight regarding the extent to which changing enumerators is effectively an issue. In particular, assuming equal quality of enumerators 1 and 3 (inexperienced) and 2 and 4 (experienced), one expects the results of column 5 to fit in between those of columns 2 and 4 if changing the enumerator is not a clear issue. Clearly, these assumptions are a bit far-fetched, in particular when applied to such a small sample of enumerators. Nevertheless, we find that column 5 results, although weak, tend to be at least as high as those of column 2, suggesting that with experienced-enough enumerators, a change of observer between two rounds of survey should not raise major issues.

Overall, results from Table 4.1 point to the importance of highly qualified enumerators. As shown by the differences between reliability coefficients, attitudinal questions seem to raise special challenges for enumerators that must be taken into account carefully. To properly assess reliability and validity issues, only enumerator 2 is considered in subsequent analyses.

Three additional tests are performed to investigate reliability issues. First, as discussed in Krueger and Schkade (2008), the reliability of attitudinal questions may greatly depend on changes of respondents' mood across survey rounds. A straightforward test is thus used to assess the significance of the relationship between changes in mood and changes in aspiration across the two rounds. Mood state was assessed by asking respondents which of five states, ranging from feeling very depressed to feeling very good, they felt most appropriate to describe their mood on that particular day. Interesting to note, no respondents changed their mood from one extreme to the other during the two weeks that elapsed between the two surveys, but 56 percent suffered from a degradation in mood and 18 percent saw an increase. Yet as is shown in Table 4.2, such a change in mood is mostly uncorrelated with eventual changes in aspiration levels across the two survey rounds.

Table 4.2—Reliability 2: Mood dependence

	Change in Income Aspiration (1)	Change in Wealth Aspiration (2)	Change in Educa- tion Aspiration (3)	Change in Social Status Aspiration (4)
Change in mood between rounds 1 and 2	411.56 (1067.48)	427.44 (6632.67)	0.14 (0.21)	0.64 (2.36)
	49	48	50	49

Source: Authors' calculations.

Second, anchoring issues relate to situations in which respondents' answers are driven in part by the responses they provided to previous questions (Kahneman and Tversky 1974). Here again, within the context of attitudinal questions, these issues are likely to be exacerbated. To test for such an issue, we randomly varied the ordering of the questionnaire across individuals. Accordingly, half of the respondents were first asked about what they thought their assigned peer's aspiration was and then asked about themselves. The other half were asked about themselves before being asked about their peers. A simple assessment of anchoring effect is thus to test for the existence of a relationship between an individual's aspiration level and the answer he or she gave for his or her assigned peer when the latter was administered beforehand. Failure to find a relationship is thus indicative of a reasonably anchor issue-free measurement tool. The corresponding tests are reported in Table 4.3. Although we find some degree of correlation between an individual's aspiration and what he or she thinks his or her peer's aspirations are, no order or interaction effects can be found, suggesting that anchoring is not a significant issue in the present context.

Table 4.3—Reliability 3: Anchoring effects

	Income		Wealth		Education		Social Status	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Other's level	0.31 (0.19)	0.39 (0.25)	0.19 (0.16)	0.27 (0.37)	0.57 (0.06)***	0.49 (0.09)***	0.27 (0.10)**	0.18 (0.14)
Questions about peer first		8913.22 (8786.19)		-8342.29 (21650.0)		-1.13 (1.72)		-9.40 (13.65)
Peer's Level × Question about Peer First		-0.11 (0.41)		-0.09 (0.42)		0.16 (0.13)		0.18 (0.21)
Observations	44	44	44	44	45	45	50	50

Source: Authors' calculations.

Notes: ** $p < 0.05$; *** $p < 0.01$.

Last, we assess the extent to which responses to the proposed measures relate to the framing of the questions themselves. In fact, although we are more interested in collecting unbounded continuous indicators based on actual values, Likert-type scales have proved consistent to generate comparable assessment across individuals. Thus, each individual was asked to place where he or she currently is and where he or she wishes to be on scales ranging from 1 to 10, with 10 being the highest corresponding level. Income and assets were not differentiated in these scales, and only a question of material well-being was asked. Questions about education and respect from the community were asked separately. The corresponding correlations with the aspiration components are presented in Table 4.4. The overall picture is one of strong positive correlation, further supporting the reliability of the instruments. Only the asset aspiration correlates poorly to the general material well-being component, which is likely related to ambiguity in the wording of the material well-being question.

Table 4.4—Reliability 4: Wording

	Income Aspiration (1)	Wealth Aspiration (2)	Education Aspiration (3)	Social Status Aspiration (4)
Likert-type scale- based measure	6457.68 (3885.56)*	8645.07 (10748.0)	1.52 (0.12)***	13.27 (2.39)***
	49	48	50	49

Source: Authors' calculations.

Notes: * $p < 0.10$; *** $p < 0.01$.

These results clearly support the overall reliability of the individual component of the aspiration measurement tool proposed. Yet despite visual scales and the various preliminary questions asked, a strong caveat is the particular

importance of reliable enumerators to administer such types of questionnaires, over and beyond what may be recommended for non-attitudinal questions.

4.3. Validity

We assess the extent to which the aspiration instrument defined here is a valid measure of aspirations per se. To this end, we devise a series of tests to assess the construct validity of the measure proposed—that is, to what extent does it correlate with expected determinants, and to what extent does it predict future-oriented behavior? To start, as emphasized by Ray (2006), one’s aspirations are influenced by the information revealed by “relevant others.” In other words, it is not information per se that matters but who it came from and whether that person belongs to one’s aspiration window. We check for the validity of our indicator, using the education component as an example. Specifically, half of the respondents were randomly chosen to be provided with the following information on returns to education after completion of the first round of survey: “Recent studies show that the returns to education are high in Ethiopia. These studies clearly established that individuals with more education achieve higher income levels compared to those with less education. For example, a person who completed secondary education obtains more than twice as much income as a person who finished primary school. Similarly, someone with primary education earns more income than someone with no education.”

A simple test for correlation between education aspiration in round 2 and whether one had received such information after round 1 shows no statistical relationship and provides initial evidence of the validity of the aspiration components collected.

Testing for validity, however, is most appropriately done using the aggregate index proposed in the Definition of an Instrument section. In fact, one is to assess whether our measure of an individual’s general propensity to aspire to a better future positively relates to expected determinants and implied behavioral outcomes. Given the reliance of Beaman et al. (2012) on a non-weighted index, we should take a moment here to discuss whether these weights may affect the indicator. First, recall from Table 4.1 that the weights obtained by experienced enumerators reached a satisfactory level of reliability. One now wishes to know whether these weights are likely to produce any change in the aggregate index—that is, whether they entail sufficient heterogeneity. Table 4.5 presents simple descriptive statistics of the weights obtained for the first round of surveys. Clearly, income and wealth seem to receive higher weights than education and social status. However, one also notes that the combination of income and wealth never reach the total number of weights possible—the maximum weights for their combination reaches 17 out of a possible 20—supporting the need to account for the multidimensional aspect of aspirations. Finally, we note a high level of heterogeneity within each component as shown by the level of standard deviation and the range between the minimum and the maximum reported. Overall, these results are clear indications of various perceptions of what individuals consider a satisfactory life. Failure to account for this could lead one to assign low or high aspiration to individuals who exert different preferences over life items.⁹

Table 4.5—Weights

	Mean (1)	Standard Deviation (2)	Minimum (3)	Maximum (4)
Income	7.08	2.04	3	11
Wealth	5.44	1.56	2	8
Education	3.57	3.52	0	12
Social status	3.60	1.32	1	7

Source: Authors’ calculations.

We now turn to the determinants of aspirations and assess how they correlate with the proposed index. Recall that from the sampling procedure, individuals were grouped into five education groups and six age groups. Taken together and applied to our sample, this led to 19 possible combinations of the three variables. It is worth noting, as a first approximation, that these groups explain 38 percent and 44 percent of the variance in aspiration at rounds 1 and 2, respectively. Further, Table 4.6 presents a series of correlations between one’s aspiration and the various classes for each of these components.

⁹ One may wish to assess the correlates of various weighting schemes. In the absence of theoretical background, such an exercise is at best exploratory and is thus not reported here. For the sake of interest, however, one might note that weights for income aspirations are positively related with age, wealth aspiration is positively related with age and negatively correlated with education, education aspiration is positively related with education and negatively correlated with age, and social status aspiration is positively related with education. Interesting to note, no effect was uncovered for gender.

Table 4.6—Determinants of aspirations

	Aspiration Index				
	(1)	(2)	(3)	(4)	(5)
<i>Education (excluding no education)</i>					
Literacy training	0.59 (0.31)*				
Primary—up to fourth	0.43 (0.41)				
Primary—fifth and sixth	0.49 (0.31)				
Junior secondary—seventh and eighth	0.64 (0.33)*				
<i>Any education? (excluding no education)</i>					
Yes		0.55 (0.26)**			0.57 (0.30)**
<i>Gender (excluding female)</i>					
Male			0.46 (0.19)**		0.53 (0.21)**
<i>Age (excluding 15–20)</i>					
20–25				0.33 (0.28)	0.63 (0.27)**
25–35				0.07 (0.32)	0.05 (0.29)
35–45				0.26 (0.52)	0.01 (0.48)
45–55				–0.27 (0.36)	0.38 (0.92)
55–65				0.08 (0.32)	0.23 (0.30)
Observations	44	44	44	44	44

Source: Authors' calculations.

Notes: * $p < 0.10$; ** $p < 0.05$.

As expected, higher education levels relate to higher aspiration (although in an eventual feedback loop). One notes, however, that except for higher degrees, it is not the degree attained, but the capacity to read or write that seems to matter more. It is also quite relevant that male respondents aspire significantly higher than their female counterparts, which is in line with more restricted opportunities for women in rural Ethiopia. Finally, and quite surprisingly, no clear relationship is found for age. Only in column 5, where the effects of all three characteristics are jointly evaluated, does one find a higher degree of aspiration for those individuals within the 20 to 25 years of age category. Taken together, these results support basic assumptions about the determinants of aspirations, namely, that aspirations respond to opportunities (whether related to gender issues or to education) but may exert a life cycle effect, with younger individuals aspiring more than their elder counterparts. Thus, the index proposed appears satisfactory for fitting with basic determinants of aspirations. Also note that similar results were obtained for the non-weighted aspiration indicator.

Although our purpose here is not to provide a thorough test of the aspiration failure framework derived from Appadurai (2006) and Ray (2006), one can go a step further in assessing the construct validity of the proposed index. In particular, both authors emphasize the importance of one's aspiration window to determine actual aspiration level, with a larger aspiration window's enhancing an individual's visualization of the steps necessary to bettering his or her future, through observations of a larger set of peers' behaviors and outcomes. An indirect test of this proposition is to relate aspirations to the frequency with which one relates to individuals outside of one's community.¹⁰ In fact, our data show statistically significant positive differences in aspiration levels between those who regularly travel to the closest city (Debre Berhan) and those who do not.

¹⁰ Clearly the relationship may run both ways, as one who aspires to more also may be more curious and may attempt to collect information more regularly.

Last, we investigate the capacity of the proposed aspiration index to predict future-oriented behavior. Our data do not allow us to directly relate aspirations to actual behavior. Thus, we use responses to three hypothetical questions related to one's potential demand for credit. The questions were framed as follows:

A banker came to you and offered to lend you any amount of money you ask ...¹¹

1. How much would you ask for if the loan were payable in 1 year?
2. How much would you ask for if the loan were payable in 5 years?
3. How much would you ask for if the loan were payable in 10 years?

Accordingly, an individual with higher aspiration should be willing to borrow more; the difference should be particularly striking for those loans allowing long-term investments. Results are reported in Table 4.7, which clearly shows the aspiration index's capacity to predict low future-oriented behavior. The results are even more consistent in that the difference between low and high aspirations regarding credit demand systematically increases with the loan's capacity to finance long-term investments. For comparison purposes, columns 4 through 6 report similar estimates using the non-weighted aspiration index. Although one notes a similar relationship for the weighted index, the magnitude and precision of the estimates are substantially lower, further supporting the need for such weighting factors.

Table 4.7—Aspiration-related behavior

	Weighted Index			Unweighted Index		
	Demand for 1-year Loan (1)	Demand for 5-year Loan (2)	Demand for 10-year Loan (3)	Demand for 1-year Loan (4)	Demand for 5-year Loan (5)	Demand for 10-year Loan (6)
Aspiration	568.58 (253.47)**	1010.69 (526.23)*	2069.80 (871.92)**	86.55 (59.01)	164.06 (121.21)	375.74 (199.99)*
Observations	44	44	44	49	49	49

Source: Authors' calculations.

Notes: * $p < 0.10$; ** $p < 0.05$.

Overall, following usability and reliability assessments, the various validity tests performed on the proposed aspiration index give significant support to its capacity to effectively measure aspirations.

5. CONCLUSIONS

Testing the promising aspiration-related theories that have been developed during the past few years requires that adequate measurement instruments be designed. Although such efforts are common in the field of psychometrics, this is less the case for the field of economics, where many studies use proxy indicators chosen ex post among a set of potentially relevant variables to account for such intangible characteristics as one's aspirations. In this paper, we attempt to provide a usable, valid, and reliable instrument to measure aspirations within standard surveys.

The designed instrument rests on an empirical definition of *aspiration* closely linked to the theoretical descriptions provided in the literature. A specifically designed individual survey was used to implement the instrument. Both the questionnaire and the sampling design were framed to facilitate tests of the questionnaire's usability, validity, and reliability. Overall, the results displayed strong support for the instrument, provided that refinements are added in future uses. The most important caveat found is the necessity of experienced enumerators who are able to adequately probe respondents.

This paper forms part of a broader effort to test aspiration-related theories in the context of Ethiopia. A first paper provides descriptive statistics showing a number of expected correlations between aspirations, aspiration windows, and aspiration failures as suggested by theory, although it does not account for major endogeneity issues pervasive in such exercises (Bernard, Dercon, and Taffesse 2011). The present paper designed an instrument to effectively measure aspirations within standard surveys. Using both papers, an experimental study is in preparation using a set of relevant documentaries on Ethiopian success stories, also made as part of this project.

¹¹ This formulation of the credit demand question was intended to make credit market imperfections in principle inoperative.

REFERENCES

- Angrist, J. D., and A. B. Krueger. 1999. "Empirical Strategies in Labor Economics." In *Handbook of Labor Economics*. Vol. 3, part A, edited by O. C. Ashenfelter and D. Card, 1277–1366. Amsterdam: North-Holland.
- Appadurai, A. 2006. "The Capacity to Aspire: Culture and the Terms of Recognition." In *Understanding Poverty*, edited by A. V. Banerjee, R. Bénabou, and D. Mookherjee. Oxford, UK: Oxford University Press.
- Beaman, L., E. Duflo, R. Pande, and P. Topalova. 2012. "Female Leadership Raises Aspirations and Educational Attainment for Girls: A Policy Experiment in India." *Science Magazine*, February.
- Bernard, T., S. Dercon, and A. S. Taffesse. 2011. *Beyond Fatalism—An Empirical Exploration of Self-efficacy and Aspirations Failure in Ethiopia*. Working Paper CSAE WPS/2011-03. Oxford, UK: University of Oxford, Centre for Studies on African Economies.
- Bertrand, M., and S. Mullainathan. 2001. "Do People Mean What They Say? Implications for Subjective Survey Data." *American Economic Review* 91 (2): 67–72.
- Blanton, H., and J. Jaccard. 2006. "Arbitrary Metrics in Psychology." *American Psychologist* 61 (1): 27–41.
- Bowles, S. 2004. *Microeconomics—Behavior, Institutions, and Evolution*. New York; Princeton, NJ: Russell Sage Foundation, Princeton University Press.
- Delavalande, A., X. Giné, and D. McKenzie. 2011. "Measuring Subjective Expectations in Developing Countries: A Critical Review and New Evidence." *Journal of Development Economics* 94 (2): 151–163.
- DeVon, H. A., M. E. Block, P. Moyle-Wright, D. M. Ernst, S. J. Hayden, D. J. Lazzara, S. M. Savoy, and E. Kostas-Polston. 2007. "A Psychometric Toolbox for Testing Validity and Reliability." *Journal of Nursing Scholarship* 39 (2): 155–164.
- Foster, J., J. Greer, and E. Thorbecke. 1984. "A Class of Decomposable Poverty Measures." *Econometrica* 52 (3): 761–766.
- . 2008. "The Foster-Greer-Thorbecke (FGT) Poverty Measures: Twenty Five Years Later." *Journal of Economic Inequality* 8 (4): 491–524.
- Greif, A. 2005. *Institutions and the Path to the Modern Economy: Lessons from Medieval Trade*. Cambridge, UK: Cambridge University Press.
- Hand, D. J. 1996. "Statistics and the Theory of Measurement." *Journal of the Royal Statistical Society, Series A (Statistics in Society)* 159 (3): 445–492.
- Hausmann, R., and D. Rodrik. 2003. "Economic Development as Self-discovery." *Journal of Development Economics* 72 (2): 603–633.
- Kahneman, D., and A. Tversky. 1974. "Prospect Theory, and Analysis of Decision under Risk." *Econometrica* 47 (2): 263–292.
- Knight, J., and R. Gunatilaka. 2012. "Income, Aspirations and the Hedonic Treadmill in a Poor Society." *Journal of Economic Behavior and Organizations* 82 (1): 67–81.
- Krueger, A. B., and D. A. Schkade. 2008. "The Reliability of Subjective Well-being Measures." *Journal of Public Economics* 92 (8/9): 1833–1845.
- Macours, K., and R. Vakis. 2009. *Changing Households' Investments and Aspirations through Social Interactions: Evidence from a Randomized Transfer Program in a Low-income Country*. Policy Research Working Paper 5137. Washington, DC: World Bank.
- Manski, C. F. 2004. "Measuring Expectations." *Econometrica* 72 (5): 1329–1376.
- Morgan, G., and M. Henrion. 1990. *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*. New York: Cambridge University Press.
- Oxford English Dictionary*, 2nd ed. 1989. <http://dictionary.oed.com/>.
- Ray, D. 2006. "Aspirations, Poverty, and Economic Change." In *Understanding Poverty*, edited by A. V. Banerjee, R. Bénabou, and D. Mookherjee. Oxford, UK: Oxford University Press.

Stutzer, A. 2004. "The Role of Income Aspirations on Individual Happiness." *Journal of Economic Behavior and Organizations* 54 (1): 89–109.

Weiner, J. 2007. "Measurement: Reliability and Validity Measures." Mimeo, Bloomberg School of Public Health, Johns Hopkins University, Baltimore.

About ESSP II

The Ethiopia Strategy Support Program II is an initiative to strengthen evidence-based policymaking in Ethiopia in the areas of rural and agricultural development. Facilitated by the International Food Policy Research Institute (IFPRI), ESSP II works closely with the government of Ethiopia, the Ethiopian Development Research Institute (EDRI), and other development partners to provide information relevant for the design and implementation of Ethiopia's agricultural and rural development strategies. For more information, see <http://www.ifpri.org/book-757/ourwork/program/ethiopia-strategy-support-program> or <http://essp.ifpri.info/> or <http://www.edri.org.et/>.

About these working papers

The ESSP Working Papers contain preliminary material and research results from IFPRI and/or its partners in Ethiopia. The papers are not subject to a formal peer review. They are circulated in order to stimulate discussion and critical comment. The opinions are those of the authors and do not necessarily reflect those of their home institutions or supporting organizations.

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE

2033 K Street, NW | Washington, DC 20006-1002 USA | T: +1.202.862.5600 | F: +1.202.457.4439 | Skype: ifprihomeoffice | ifpri@cgiar.org | www.ifpri.org

IFPRI-ESSP II ADDIS ABABA

P.O. Box 5689, Addis Ababa, Ethiopia | T: +251.11.617.2000 | F: +251.11.646.2318 | mahlet.mekuria@cgiar.org | <http://essp.ifpri.info>

ETHIOPIAN DEVELOPMENT RESEARCH INSTITUTE

P.O. Box 2479, Addis Ababa, Ethiopia | T: +251.11.550.6066 ; +251.11.553.8633 | F: +251.11.550.5588 | exe-director@edri.org.et | www.edri-eth.org

This Working Paper has been prepared as an ESSP II output, and has not been peer reviewed. Any opinions stated herein are those of the author(s) and do not necessarily reflect the policies or opinions of IFPRI.

Copyright © 2012, International Food Policy Research Institute. All rights reserved. To obtain permission to republish, contact ifpri-copyright@cgiar.org.