The World Bank Economic Review, 31(3), 2017, 670–686
doi: 10.1093/wber/lhw026
Advance Access Publication Date: April 19, 2016
Article

Identifying Gazelles: Expert Panels vs. Surveys as a Means to Identify Firms with Rapid Growth Potential

Marcel Fafchamps and Christopher Woodruff

Abstract

We conduct a business plan competition to test whether survey instruments or panel judges are able to identify the fastest growing firms. Participants submitted six- to eight-page business plans and defended them before a three- or four-judge panel. We surveyed applicants shortly after they applied and one and two years after the competition. We use follow-up surveys to construct measures of enterprise growth and baseline surveys and panel scores to construct measures of enterprise growth potential. We find that a measure of ability correlates strongly with future growth, but that the panel scores add to predictive power even after controlling for ability and other survey variables. The survey questions have more power to explain the variance in growth. Participants presenting before the panel were given a chance to win customized management training. Fourteen months after the training, we find no positive effect of the training on growth of the business.

JEL classification: L26, O12, O31

A majority of the labor force in many Sub-Saharan African countries is self-employed. Growth of wage employment is a key goal for many policy makers in the region. Wage jobs sometimes are created in large number by a single, large firm but more often are created a handful at a time by the modest expansion of large numbers of small firms. However, only a minority of microenterprises ever hire any employees. Most microentrepreneurs in low-income countries do not aspire to grow, and indeed, may not be able to manage larger enterprises. Can we identify the minority of small firms with the potential to create employment? The payoff from doing so is that policy makers and NGOs targeting job creation would be able to develop policies and programs to stimulate more rapid expansion among these firms.

In this paper, we report on a field exercise comparing two methods of identifying small firms with the potential for rapid growth—commonly referred to as "gazelles." We conducted a business plan competition in Ghana, with panels of judges comprised of successful business owners—who themselves started very small businesses—consultants and other experts on small enterprises in Ghana. The panels judged written business plans and oral presentations by entrepreneurs with between two and 20 employees. Prior to the competition, each applicant had the opportunity to attend a three-day class that provided

Marcel Fafchamps is a senior fellow at the Freeman Spogli Institute for International Studies, Stanford University; his email address is fafchamp@stanford.edu; Christopher Woodruff (corresponding author) is a professor in the Department of Economics, University of Warwick, Coventry UK CV4 7AL; his email address is c.woodruff@warwick.ac.uk. We thank the World Bank's Multi Donor Trust Fund and the Social Protection and Labor Division and the Templeton Foundation for financial support. IPA-Ghana and Caroline Kousssiaman provided exceptional assistance on the logistics of the project. We also thank Afra Kahn and Paolo Falco for assistance. A supplemental appendix to this article is available at: https://academic.oup.com/wber.

assistance in preparing the business plan. We used surveys conducted before the entrepreneurs attended the business plan writing course to gather information on the entrepreneurs. The survey responses provided the second—and much cheaper—means of identifying fast-growing firms and an assessment method with which to compare the panel judgments.

Rapidly growing new and existing small firms are an important source of job creation in many contexts. Using comprehensive data from the United States, Davis et al. (2007) find that only about 3 percent of Schedule C enterprises (nonemployers) ever hire a paid employee. But the vast number of Schedule C enterprises means that these expansions account for 28 percent of all new employers and 20 percent of new employment. Cabral and Mata (2003) show that among the Portuguese enterprises with one employee in 1984 that survive until 1991, half grew to more than one employee by the later date. There is limited evidence from developing countries on the dynamics of microenterprises. In a series of "Enterprise Maps" in several African countries, Sutton and coauthors study the origins of 50 leading firms in each country. In Ghana, Sutton and Kptenty (2012) find that 15 of these largest firms began as small-scale startups, suggesting that rapid growth—while not common—sometimes does occur. Closer to the exercise we have in mind in this paper, de Mel et al. (2010) use retrospective data from a cross-sectional survey of small and medium-sized employers in Sri Lanka to show that 12 percent of firms with more than five employees had no employees during their first year of operation.

Business plan competitions are an increasingly popular method of identifying fast-growing firms.[1] A typical competition has a panel of experts—successful business people, consultants, and financial analysts—who judge a combination of written business plans and oral presentations. The winners of the competition receive mentoring and, often, cash prizes. In developing countries, business plan competitions are seen as being one means of identifying firms with rapid growth.

We might ask, from a policy perspective, why should we be concerned with an enterprise's potential for growth? The focus in rapid growth comes at least in part from an interest in job creation among policy makers. But perhaps policy makers should instead be concerned with identifying enterprises that face constraints to growth that policy interventions are best able to overcome. In this regard, in our sample we suggest that there is substantial overlap between those facing constraints and those that have the most potential for growth. Our competition is held among enterprises that have been in business for an average of nine years, with three-quarters of them at least three years old. A minority (25 percent) have ever had a bank loan or had formal entrepreneurship training (42 percent). Those identified by panelists as having potential for growth are apparently held away from their optimal size by some constraint, which could be capital, entrepreneurial skills, or simply a lack of self-belief by the enterprise owner. We are unable to tackle all of these potential constraints in a single project and so focus on entrepreneurial skills as a potential constraint. As businesses grow in size, formalization of processes and skills in managing employees become more important. We use training and consulting as a prize in the competition, randomizing the probability of winning the training based on the quartile of the ranking by the panel. This allows us to test whether training is a key constraint among those viewed by the panel as being farthest from their optimal size. Unfortunately, as we will see, we are not able to provide a clear answer to the question of whether those identified as having the highest potential for growth benefit more (or less) from training relative to other firms.[2]

1 Technoserve runs an annual business plan competition in Ghana, which it calls Believe, Begin, Become. The main target group of that competition is white collar workers not currently self employed. See Klinger and Schundeln (2011) for analysis of a similar competition run by Technoserve in three Central American countries. McKenzie (2015) reports on the results of a business plan competition in Nigeria while Fafchamps and Quinn (2015) report on a business plan competition in Ethiopia, Tanzania and Zambia.

2 See McKenzie (2015) and Fafchamps and Quinn (2015) for business plan competitions that instead relax the capital constraint. We discuss the comparison in results in more detail below.

A second key concern is whether we should expect the panels of "experts" to be effective in identifying entrepreneurs with the greatest potential. Here, both active debate and some convergence emerge from the literature on the question of whether and when expert opinion is likely to be helpful (Shanteau 1992; Kahneman and Klein 2009). Shanteau assesses the predictive value of experts by listing characteristics of good and bad decisions. Kahneman and Klein summarize three conditions that, in combination, increase the accuracy of expert forecasts: (1) the outcomes being judged are reasonably predictable; (2) the experts have extensive experience making those judgments; and (3) the experts receive rapid and continuous feedback on the accuracy of their initial judgments. When these conditions are not met, they conclude that simple linear combinations of traits are likely to be more predictive.

By these criteria, we might expect a relatively poor performance from the panelists. At a minimum, the relevant outcome—growth of the enterprises—is difficult to predict and subject to many factors outside the control of the entrepreneurs. Whether the experts make these judgments regularly and whether they receive feedback on the accuracy of their judgments is not clear. Regular feedback on past judgments may be more common for the consultants than for the successful business people on the panels. The latter are likely to make systematic judgments of entrepreneurial outcomes and to receive regular, repeated feedback on those judgments only if they are involved in angel finance.

A further element in the present context is the extreme level of heterogeneity in ability and aspirations of small-scale business owners in lower-income economies such as Ghana. The business plan competition panels may be able to differentiate the prospects of the upper and lower tail from the majority in the middle, but they may not be able to differentiate within the middle group.

Considering both the lessons from the literature on expert decision making and the context, we see the question of whether microenterprises with potential for growth can be identified as empirical in nature. With this in mind, we test the ability of both panelists and surveys to predict enterprise growth. We measure growth using follow-up surveys conducted with the participating entrepreneurs roughly one and two years after the panel meetings. These follow-up surveys give us outcomes on growth over the medium run. We use the baseline surveys, panel rankings, and follow-up surveys to assess whether the expert panels are able to predict which enterprises grow, and if so, whether they are better predictors of growth than simple measures from survey questions.

The psychology literature commonly pits experts against algorithms. We begin by comparing the predictions of the panel with predictions based on variables from the survey questionnaire. One problem inherent in this exercise is that while the panel rankings are summarized in one or two measures, the survey contains a very large number of questions and, almost certainly, some of them will correlate with the subsequent growth of the enterprises. We begin by limiting ourselves to a set of core measures which we think are likely to be viewed as at least potentially important—measures of the ability of the owner, past borrowing activity, and management practices.

Using any of several measures of growth, we find that the ability measure from the survey—a combination of nonverbal reasoning tests, a numeracy test, years of formal schooling and financial literacy—predicts growth. There is also some evidence that baseline management practices are associated with future growth. We also find that the entrepreneurs that panels rank higher grow faster. But the survey responses explain a larger share of the variance in outcomes, regardless of whether we use changes in employment, revenues, or profits as the measure of growth. However, even though by this analysis of variance measure the surveys best the panels in a "horse race," we also need to ask whether the surveys and panels combined give us better predictions than the surveys alone. That is, do the panels add predictive power on top of the surveys—and enough predictive power to justify their use? We answer the first part of this question, at least, in the affirmative. Even after controlling for the survey measures, the panel ranking remains a significant predictor of growth.

In addition to understanding the extent to which we might expect job creation from microenterprises if constraints to their growth were reduced, separating gazelles from the mass of microenterprises might be valuable for designing more appropriate SME support programs. Governments and NGOs currently direct large amounts of resources toward the SME sector.[3] We provided more customized management training to some of the enterprises entering the competition as an incentive to participate. We randomly allocated training to owners presenting before the panel, with the odds of receiving training increasing in the panel ranking, as described below. However, we find little evidence that the training stimulated growth in the short run. (The last follow-up survey was just over a year following the training.)

A brief roadmap of the paper is as follows: We begin by reviewing the design and protocol, including a discussion of the panelists and the baseline survey. We next describe the results, first on the relationship between panel rankings, survey measures, and growth, and then on the effects of training. We conclude with a brief discussion of the results and implications for the need for experts in separating microenterprises with greater potential for growth.

I. Design and Protocol

We launched a business plan competition in Accra and Tema, Ghana, in January 2010. Our initial design called for a sample of 400 applicants, of which 100 were to be randomly selected as a pure control group and 300 were to receive a three-day business plan training course. We planned to provide more extensive training for half of the 300 businesses randomized into the business plan competition, with the probability of receiving the training increasing in the score received from the panel of judges in a manner we describe below.

We announced the business plan competition through advertisements in the newspaper and on radio in the Accra-Tema metropolitan area. We also worked with the Association of Ghanaian Industries (AGI), both for the recruitment of participants and the recruitment of judges. Finally, we selected 15 neighborhoods with large concentrations of small businesses and conducted a door-to-door marketing program using local research assistants. The application form for the competition (shown in appendix S1, available at https://academic.oup.com/wber) was available at AGI, at the project office, and on the internet.

The sample design was based on a successful pilot project we carried out in Accra between January and March 2009. The pilot had the modest goal of obtaining a sample of around 20 enterprises with between three and 14 paid employees whose owners were between the ages of 20 and 40. With little marketing effort, we received 27 eligible applications for the pilot phase. Of these, 23 arrived on the first day of a two-day training course geared to writing a simple business plan. The course was designed and offered by CDC Consult Limited, a local consulting company with extensive experience working with SMEs in Ghana. The explicit goal of the course was that each participant would leave with a draft business plan. Twenty-one participants completed the full two days. These participants were then given one week to submit a three- to five- page business plan, and 19 submitted the business plan. Thus, two-thirds of the applicants in the pilot followed through to the business plan submission.

Based on the experience with the pilot, we forecasted that a target of 400 participants for the full project was attainable. However, by the initial March 1, 2010, deadline, we had received only 92 applications. We began the business plan training courses for this initial group, while also announcing a second deadline of May 1, 2010. By the latter date, we had received an additional 75 applications. However, only 143 of the combined group of 167 applicants met the application criteria. We conducted baseline surveys for these 143 applicants. Given the challenge in reaching the targeted sample size, we extended

3 To give one example, the ILO's Start and Improve Your Business training program now has around 100 million alumni in 95 countries. The participation of most of the participants in that program has been heavily subsidized.

the project to the city of Kumasi in May, again using a combination of print and radio advertising and door-to-door marketing. The response in Kumasi was much more robust. Though Kumasi is substantially smaller than the Accra–Tema metropolitan area, we receive more than 200 applications, of which we confirmed eligibility and completed the baseline survey for 192. Thus, the combined sample from Accra and Kumasi is 335, still somewhat short of the initial target in spite of higher-than-planned effort level. Given this, we adjusted the initial design to eliminate the pure control group. We offered all of the applicants the basic business plan training.[4]

The majority of our applicants came from direct marketing in targeted neighborhoods. Though we lack data with which to benchmark our sample to the population, the fact that the majority came from blanket door-to-door marketing in targeted neighborhoods suggests that we have a sample roughly representative of small businesses interested in participating in a business plan completion.[5]

Each applicant completing the baseline survey was subsequently invited to participate in a three-day business plan training course. As in the pilot, the course was offered by CDC Consult. We extended it to three days based on feedback from participants and trainers following the pilot. We initially assigned applicants to a training session starting on a specific date. However, compared with our experience in the pilot, we experienced high nonattendance rates in the first training sessions of the full-scale project. We therefore gave the participants more flexibility in selecting the training session they would attend, asking applicants to tell us which sessions were most convenient for them.

Even with additional flexibility, nonattendance rates were higher than in the pilot. Around 30 percent (100 of the 335) failed to attend any of the business plan training, and an additional 6 percent (19) failed to complete the training. Although there were more responses to the call for applications in Kumasi, dropout rates were somewhat higher there. Almost 38 percent of the Kumasi applicants did not complete the initial training, compared with around 32 percent of those from Accra-Tema; the difference in dropout rates between cities is not statistically significant. Of the 216 who completed the initial training, 152 (45 percent of the baseline sample) submitted a business plan and 141 (42 percent) presented the plan before the panel of judges. Dropout rates in the later phases were lower in Kumasi, so that similar portions of the baseline sample completed the panel presentations in the two areas (42 percent in Kumasi vs. 41 percent in Accra-Tema). We believe that dropout behavior is interesting in itself, and we examine the characteristics of those who drop out at various points below in the results section.

II. Judging the Business Plans

We organized panels of three or four successful small business owners and consultants with extensive experience working with small businesses in Ghana. The panels of four judges give some additional statistical power, which we viewed as important given that the sample size that was smaller than we had initially anticipated.[6] The majority of the judges were consultants rather than business owners, though most panels contained at least one business owner. Each panel was assigned 12–16 business plans.[7]

The judges first received the written business plans. These were read and scored according to five criteria: the description of the business concept; the definition of the market; the description of the current organization; the financial statements; and the overall organization of the business plan. The written

4 This implies that we are unable to estimate the impact of the business plan training course on the outcomes of interest.

5 The language for the business plans, panels and training was English, which is the official language in Ghana. English is widely used in schools and government administration. Thus, given our focus on identifying small firms with potential for rapid growth, we do not view this as constraining the sample.

6 In two panels in Kumasi and one in Accra, one of the judges did not show up for the presentations, so we use the scores of the remaining three judges.

7 Wrap-up panels in Accra/Tema and Kumasi received only six and 10 proposals, respectively.

plans were marked by each judge independently, before the judges met for the first time as a panel. The judges then met on two occasions—typically, once on a weekday evening and once on a Saturday—to hear presentations of the business plans by the applicants. Each applicant was allotted 30 minutes for his or her presentation, divided into a 15-minute presentation and a 15-minute question-and-answer session.

The judges were asked to give specific marks on the oral presentation for preparation, confidence, understanding of the business, the ability of the entrepreneur to make his/her case, and the ability to answer questions. Next, the three judges were asked to give overall marks to the applicants, combining the written business plans and the presentation, on five criteria: the applicant's business acumen; how well he or she runs the existing business; the strategy for growth; his or her ability to manage a growing enterprise; and his or her ability to articulate goals and vision. Finally, we added two summary scores that we use for much of the analysis. The first asks each judge to rank on a scale of 0 to 100: "Based on both the written business plan and the oral presentation, how would you rate the business' growth potential?" We refer to the score given in response to this question as the "comprehensive growth measure." The second asks for a similar 0 to 100 ranking: "Based on both the written business plan and the oral presentation, how likely would you be to recommend to an angel investor that (s)he invest in this business?" We refer to this score as the "angel investor measure."

Although all judges use the same range for scores, some individual judges and even panels may have ranked applicants more harshly than others. To make comparisons across panels more meaningful, we calculate a standardized measure for each score given by judge j to applicant i:

$$\frac{score_{ij} - score_j}{sd(score_j)},$$

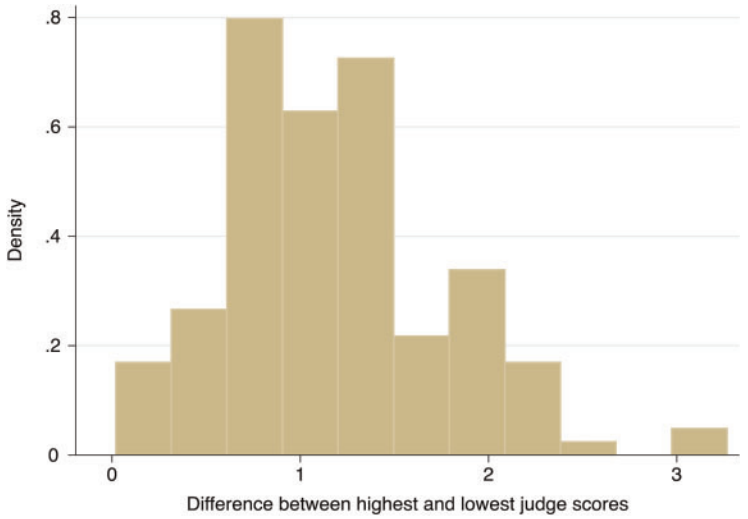
where $score_j$ is the average score given to all applicants by judge j and $sd(score_j)$ is the standard deviation of judge j's scores. We take the standard deviation of this measure across the scores given by various judges to an applicant as a measure of the disagreement about the applicant. We calculate this for three main summary measures: the overall score, which sums up the written and oral presentation scores; the comprehensive growth score; and the angel investor measure.

The first issue we investigate is whether the judges agree on the rating and ranking of the applicants they reviewed. Fig. 1 shows the extent of disagreement on individual proposals, using the aggregate measure of prospects for growth. For ease of interpretation, in the graph we take the difference between the standardized ranking of the most favorable and least favorable judge. We see considerable disagreement among panel members. The median difference is just under 1.2 standard deviations. This would arise, for example, if the most pessimistic judge rates the potential for growth 0.6 standard deviations below his mean ranking, while the most optimistic judge rates the growth prospects 0.6 standard deviations above his mean ranking. For a cluster of entrepreneurs, this gap exceeds two standard deviations, indicating considerable disagreement. We know of no absolute standard against which to measure this dispersion, but a high-low range of less than a standard deviation strikes us as substantial agreement. The panel judgments for around 42 percent of the entrepreneurs are below this level of disagreement.

III. Results

Ultimately, we are interested in two research questions. First, we want to know whether survey responses can predict the growth potential of firms. This question is of particular interest because business plan competitions are expensive to run. Second, we want to know whether the panels can identify entrepreneurs with more growth potential in a way that adds to the predictive power of the survey responses. Though we have noted that we are unable to precisely place our sample in the population of microenterprises, we are able to say something about selection from the initial application—and baseline

Figure 1. Within Applicant Difference between the Highest and Lowest Standardized Ranking of Judges on the Panel.



Source: Authors' calculations based on survey data.

survey—to the point of presenting before the panel. We begin our discussion of the results by examining the patterns of attrition within the project timeframe.

Analyzing Dropouts

Our analysis is conducted with a sample of owners who wish to participate in a business plan competition and are able to complete all the steps required to do so. As we noted above, we lack data to compare this sample to the population of small firms in Ghana. However, the availability of baseline data for the full sample of applicants allows us to examine the characteristics of participants who drop out between the application and the presentation.

Table 1 shows summary statistics for the sample of 335 applicants divided into three groups: the 119 who applied but did not complete the business plan training course; 76 who completed the business plan training but did not present a business plan before the panel; and 140 who presented before the panel. The data in the table come from the baseline survey, which was completed by all 335 applicants. The *p*-values in the last column compare the means of early dropouts with the means of those presenting before the panel. There are two fairly clear patterns in the attrition. First, there is an indication that opportunity costs matter. Those who drop out early work longer hours in a normal week (56 vs. 52 hours) and are more likely to be in the trade sector (56 percent vs. 36 percent). Second, by various measures, the early dropouts appear to be from the left-hand tail of the entrepreneurial skill distribution. Presenters have more schooling (14.5 vs. 13.4 years), have slightly higher Raven nonverbal reasoning scores and score better on a four-question test of financial literacy.[8] Those presenting are also more likely to use a computer in the business, have slightly better management practices, and are more willing to take risks. Finally, presenters are more likely to be female, which may be correlated with either opportunity cost or business potential. The difference in the baseline size of the business does not significantly predict attrition.

In contrast, there is almost no difference between those who completed the training without presenting before the panel with those who presented before the panel (column 2 vs. column 3). The differential

8 See appendix II for a description of the survey questions used for specific measures.

Table 1. Baseline Characteristics of the Sample by Treatment Group

	Applicants not completing business plan training	Completed business plan training, but did not present	Presented before panel	T-stat, column 3 vs. 1
Variables related to entrepreneur				
Age of owner (years)	38.9	39.6	39.6	0.53
Owner is female	15.1%	26.3%	26.4%	*0.03*
Owner ever married	75.6%	81.6%	71.4%	0.45
Years of schooling	13.4	14.2	14.5	**0.01**
Raven non-verbal score	2.8	3.4	3.3	*0.08*
Digtspan recall score	6.4	6.9	6.3	0.60
Counting backwards by 7	60.5%	50.0%	56.0%	0.44
Financial literacy	1.7	2.1	2.0	**0.01**
Self-assessed English ability (1=Beginner; 3=Fluent)	2.5	2.6	2.6	*0.04*
Previous bank loan	29.9%	34.2%	40.0%	*0.09*
Variables related to business				
Retail or wholesale trade sector	56.3%	38.2%	36.4%	**0.01**
Manufacturing sector	15.1%	25.0%	30.0%	**0.01**
Number of paid empployees	4.3	4.0	4.2	0.91
Capital Stock	GHS 14,497	GHS 10,874	GHS 14,859	0.94
Profits (truncated 99th percentile)	GHS 1,652	GHS 1,967	GHS 1,371	0.43
Normal hours worked per week	56.7	55.8	52.2	*0.03*
Age of firm (years)	9.6	9.0	9.1	0.56
Percentage of equipment which is manual	57.7%	51.4%	58.0%	0.94
Registered for taxes	63.2%	68.4%	55.0%	0.18
Uses computer in business	38.5%	52.6%	51.4%	*0.04*
Variables related to management practices				
Overall management practices score (0–27)	12.4	15.2	13.9	*0.06*
Marketing practices score (0–9)	3.6	4.3	4.0	0.24
Stock control score (0–4)	2.0	2.3	2.3	*0.06*
Book-keeping score (0–10)	5.5	6.6	6.1	0.16
Finanaical score (0–4)	1.3	1.9	1.7	*0.03*
Variables related to attitudes of owner				
Willingness to take risks	5.8	6.9	6.6	*0.02*
Believes businesses fail mostly because owners are not skilled enough or do not work hard enough	62.2%	61.8%	59.3%	0.64
Expected number of employees in five years time	10.8	11.9	15.3	0.13
Optimism	3.1	2.9	3.0	0.67
"Success" means expanding or growing the business	69.7%	75.0%	74.3%	0.42
Number of Observations	119	76	140	

Source: Authors' calculations based on survey data.

attrition thus appears to come at the point of the business plan training course. In important ways—for example, size by employment, profits, and capital stock—the subsample presenting before the panel is similar to the full population of applicants, although those presenting are more sophisticated in some dimensions. The attrition appears to reduce the heterogeneity of the sample the panel is judging. However, we do not view this as a major concern for our purpose because the characteristics that are over-represented among those presenting before the panel are ones we expect to be associated with greater potential for growth.

Predicting Growth

We first explore whether survey responses predict which enterprises grow in the period between baseline and follow-up. We conducted follow-up surveys in July and August of 2011 and again in August and September 2012. The resurvey rates were 80.3 percent and 85.7 percent in the first and second follow-up, respectively. We were able to resurvey at least once all but 15 of the original 335 firms (96 percent), and 138 of the 141 firms (98 percent) that presented before the panel of judges.[9] The resurvey rates are reasonable round-to-round given the nature of the businesses and excellent in total.

We use these data to ask what baseline measures—survey questions or panel rankings—correlate with the subsequent growth of the enterprises. We use four measures as growth measures: the number of paid employees; the level of sales; the level of profits; and the level of investment in the year prior to the follow-up survey.[10] In all four cases, we take hyperbolic sine transformations to address concerns with long right-hand tails. Since each of the four variables provides a noisy measure of enterprise growth, we also combine them into a single measure of growth by standardizing and then averaging them. We first look at outcomes using the four individual measures, but rely on the combined measure for most of the analysis. We estimate regressions with the following specification:

$$Y_{it} = \alpha + \gamma_1 Score_i + \sum_{n=1}^{5} \beta_n Z_{ni} + \theta Y_{i0} + \delta' X_i + \epsilon_{it},$$

where *Score* is the panel score, the Z_i are characteristics measured in the survey (such as the ability, credit, and other measures for owner *i)*, Y_{i0} is the baseline measure of the dependent variable, and the X_i are other controls. We use the hyperbolic sine transformations for the baseline measures as well.[11]

Given the modest sample size of 140 entrepreneurs presenting before the panel and the large number of measures on which we have baseline survey data, given time we could almost certainly find significant predictors of growth in the survey data. To avoid a pure data mining exercise, we limit the survey data to five categories of information: a measure of ability; two measures of attitudes; management practices; and access to credit, including previous loan experience. For ability and attitudes we use discriminant analysis to find the first principal component of answers to related survey questions.

We begin by assessing whether growth is predicted more accurately by ability or attitudes, or a combination of both. To measure ability, we combine five measures using discriminant analysis. The measures are: years of schooling; the score on a Raven nonverbal reasoning test; the maximum number of digits recalled on a digitspan recall test; successful recitation of the seventh number in the series counting backward from 100 by seven; and the score on a four-question test of financial literacy.[12] All five measures are positively correlated with one another, and all five enter the first PCA positively and with roughly equal weights.

A second measure relates to previous and current access to credit. This might be viewed as a signal of ability from formal or informal lenders. We combine responses to three questions: whether the owner has previously received a bank loan; whether the owner currently buys any inputs on credit (i.e., whether she/he receives trade credit from suppliers); and whether the owner says she/he would be able to borrow 5000 GhC from any source to invest in the business. Again we use discriminant analysis to extract a

9 There were three firms which were resurveyed in both follow-up rounds, and for which the name of the owner did not match the name of the owner in the baseline survey. We dropped these three firms from the analysis, a decision that has no material effect on the results, but was taken to be conservative.

10 Firms that had gone out of business by the time of the follow-up survey were assigned values of zero for all four outcome measures.

11 One right-hand side variable which is missing from the equation is assignment to training, which might have affected outcomes by the second follow-up round. Since assignment to the training treatment was conditional on the panel ranking, there is no simple way to control for training in this regression. We discuss this issue later in this section.

12 Selected survey questions are shown in appendix 2.

common factor. As with the ability measure, all three components enter positively and with similar weights. Third, we measure management practices using a diagnostic instrument developed in de Mel et al. (2012). This diagnostic tool asks a series of questions related to marketing, bookkeeping, stock control and financial planning. The responses, which are combined linearly on a scale of 0–29, are detailed in appendix S3.

The first four columns of table 2 show the correlation between these three ability measures and the four measures of subsequent growth. Column 5 uses the combined measure of growth. The regressions control for the baseline measure of the dependent variable. We add controls for the sector of activity (retail or manufacturing rather than services) and a dummy indicating location in Kumasi. These control for possible shifts in demand in the region/sector where the firm operates. We also add controls for characteristics of firms and owners that are expected to be correlated with growth—the age of the firm, the age of the owner (we expect younger firms and firms with younger owners to grow faster), and the gender of the owner. Finally, we include judge-panel fixed effects.[13] The results are generally not sensitive to the inclusion of the additional controls, and we later show regressions excluding all but the panel- and sector fixed effects. Among these controls, we find that panelists find significantly more potential for growth among manufacturers and less among retailers. We find no association between predicted growth and the age or gender of the owner, or the age of the firm, proxied with a dummy indicating the firm is at least five years old.

The baseline survey asked owners if they had previously participated in any business training programs. Just over 40 percent of the owners said that they had. These were almost always programs run by NGOs, but we have no further information on their content or intensity. Previous training is unrelated to most demographic and other controls, including age, gender, location, and sector. However, there is a strong positive correlation between training and ability, suggesting that perhaps the most able of the owners had previously sought out training.[14] When we include previous training in the table 2, column 5 regression (not shown), we find that previous training is positively associated with growth at the 5 percent level, but that ability remains significant at the 4 percent level.

A part of the sample received training as part of the project between baseline and the first follow-up. The treatment was designed in such a way that those ranked higher by the panels had a higher probability of receiving the training. The sample in table 2 includes both the treatment and control groups of the training intervention. An obvious concern is that training may be correlated with ability and that training may have affected growth. In that case, what we measure as ability (or later, as the panel score) may actually reflect training. As we will show later in the paper, training had no significant effect on subsequent growth. But when we run the regression in column 5 on the sample of firms not assigned to training, the results are qualitatively the same. The measured effect of ability, for example, is 0.12 rather than the 0.10 shown in column 5. The smaller control sample results in a loss of statistical power (e.g., the ability result is significant only at the .09 level). Given the modest sample size in the entire project and the fact that the results are similar in the full- and control samples, we use the full sample for the remaining regressions.

The results on table 2 show that ability and management practices are most closely associated with subsequent growth. Ability is significantly associated with growth measured by employment and profits, and management practices are significantly associated with growth measured by revenues and

13 The panel fixed effects absorb the Kumasi location variable. We include it in the list only because in some later specifications we drop the panel fixed effects but retain the location dummy.

14 Those with previous training were 24 percentage points more likely to have completed the (79 percent vs. 55 percent, $p < .01$), and 22 percent more likely to have presented before the panel (56 percent vs. 35 percent. $P < .01$), indicating that the sample is weighted toward those with previous business training.

Table 2. Predicting Growth from Direct Measures

	Employment (1)	Revenues (2)	Profits (3)	Investment (4)	Aggregate Growth (5)	Employment (6)	Revenues (7)	Profits (8)	Investment (9)	Aggregate Growth (10)
Ability score	0.14***	0.17	0.29*	0.28	0.10**	0.11*	0.12	0.22	0.40	0.08*
	(0.05)	(0.14)	(0.16)	(0.25)	(0.04)	(0.06)	(0.14)	(0.17)	(0.28)	(0.05)
Credit (0–3)	0.03	0.28	0.21	−0.07	−0.00	0.03	0.24	0.18	−0.03	−0.01
	(0.06)	(0.23)	(0.22)	(0.23)	(0.06)	(0.06)	(0.22)	(0.24)	(0.24)	(0.06)
Management score (0–29)	0.01	0.08***	0.06	0.12**	0.01	0.00	0.07**	0.05	0.13**	0.01
	(0.01)	(0.03)	(0.04)	(0.05)	(0.01)	(0.01)	(0.03)	(0.05)	(0.05)	(0.01)
Attitudes (Growth)						0.08	0.25	0.21	−0.39	0.05
						(0.07)	(0.18)	(0.25)	(0.27)	(0.05)
Attitudes (Control)						−0.09	0.21	−0.09	0.15	−0.01
						(0.06)	(0.16)	(0.20)	(0.22)	(0.05)
Observations	229	224	221	229	229	229	224	221	229	229
R-squared	0.464	0.231	0.140	0.201	0.361	0.476	0.246	0.144	0.209	0.364

Notes: Robust standard errors in parentheses clustered at the firm level, ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

The sample includes only owners who presented before the panel of judges. The dependent variables are levels in the follow-up surveys and all regressions control for the baseline value of the dependent variable. The dependent variables in columns 1–4 and 5–9 are all transformed using a hyperbolic sine transformation. Revenue is the average of the two months prior to the survey; investment is capital purchased in the year prior to the survey. Columns 5 and 10 are the sum of the standardized values of the 4 growth measures. All regressions include controls for owner age and female, manufacturer, retail, firm <5 years old, location in Kumasi, panel and wave 1 fixed effects.

Source: Authors' calculations based on survey data.

investment. Ability is also significantly associated with the aggregate measure of growth, but management practices and access to credit are not.

Next we consider attitudinal measures of the entrepreneurs collected in the baseline survey. These are arguably softer measures, and an area where we expect the panelists to have an advantage in discerning differences. We construct two measures of attitudes, which we refer to as attitudes toward growth and attitudes toward control. The construction of these measures is described in more detail in appendix S3, but the growth measure incorporates direct questions about aspirations to grow and measures of risk attitudes. The control measure combines measures of optimism, trust, and internal locus of control.

Neither attitudinal measure is significantly associated with subsequent growth, regardless of which growth measure we use. This may reflect a lack of association between these attitudes and growth or inherent difficulties in measuring attitudes. The predictive coefficients of "harder" measures of ability and practices are not much affected by the inclusion of the attitudes variables, though the ability measure loses significance in the employment growth regression.

Recall that approximately half of the applicants dropped out before reaching the business plan competition phase of the project. We have baseline information for essentially all of the dropouts, and at least one follow-up survey for all but 19 of the dropouts. In table 3, we use the full sample data to examine differences in growth rates between those who completed the competition and those who dropped out. We expect that the selection process at work in our experiment is similar to the selection process in a typical business plan competition, that is, it yields a sample of entrepreneurs willing and able to present a business plan in front of panel of judges. Table 3 thus provides information about those micro and small enterprises that are likely to drop out of a business competition.

Table 3. Predicting Growth from Direct Measures

	Aggregate Growth (1)	Aggregate Growth (2)
Ability score	0.04**	0.06**
	(0.02)	(0.02)
Ability * dropped out		−0.04
		(0.03)
Attitude Score	0.01	0.03
	(0.02)	(0.03)
Attitude * dropped out		−0.03
		(0.04)
Dropped out	0.05	0.05
	(0.07)	(0.07)
Observations	506	506
R-squared	0.156	0.161

Notes: Robust standard errors in parentheses clustered at the firm level, ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

The dependent variable is the sum of the standardized values of the 4 growth measures. The dependent variables are levels in the follow-up surveys and all regressions control for the baseline value of the dependent variable. All regressions include controls for owner age and female, manufacturer, retail, firm <5 years old, location in Kumasi, panel and wave 1 fixed effects.

Source: Authors' calculations based on survey data.

Business plan competitions would not be a means of identifying fast-growing firms if we find that the dropouts grow faster than those remaining in the competition. The first column of table 3 indicates instead that dropouts grow neither slower nor faster than those remaining in the competition—though we do find a small positive, nonsignificant, coefficient. In the second column of table 3, we interact the dropout variable with the ability and attitudes measures, to test if the relationship between the survey

measures and growth differs in the two subsamples. Again we find no significant differences, though both interactions are negative, indicating that the survey is, if anything, less able to determine which firms grew among the dropouts. These results are somewhat reassuring from the perspective of using business plan competitions to identify fast-growing firms, because they suggest that the fastest-growing firms did not disproportionately drop out of the competition.[15]

Do Panels Add Predictive Power?

The panelists rated both the written business plan, the oral presentation, and the overall prospects for each entrepreneur. We then asked for two summary measures, one on the prospects for growth and the other asking how attractive the enterprise would be for an angel investor. The two overall measures have a correlation of over 0.95, while the overall written plan score correlates with the aggregate growth score at the 0.47 level. There is also a significant positive—but much lower—correlation between the overall growth score and the survey measures of ability (0.24), credit (0.25), management practices (0.23), and attitudes to growth (0.37).

With this in mind, we ask whether panels can improve on the predictive power of the survey questions using the overall growth score as the panel measure. We proceed by adding the panel measure to the regression and seeing whether it enters significantly—even when we control for the survey measures—and how much more of the variance in growth the regression explains. We report the results from this exercise on table 4, using several specifications of the survey measures. The sample is limited to those who presented before the panels. The first column of table 4 repeats the regression from column 5 in table 2, for comparison. In column 2, we include the aggregate score from the panel without the survey measures. The score is highly significant, and indicates that a one standard deviation difference in panel score is associated with a 0.16 standard deviation difference in subsequent growth. By itself, however, the panel score does not explain as much of the variance (0.337) in growth as the three ability measures constructed from survey responses (0.361).

Table 4. Predicting Growth from Direct Measures

	Aggregate Growth (1)	Aggregate Growth (2)	Aggregate Growth (3)	Aggregate Growth (4)	Aggregate Growth (5)
Panel's Comprehensive Growth Score		0.16***	0.12**	0.12**	0.10*
		(0.05)	(0.06)	(0.05)	(0.06)
Ability Score	0.10**		0.08**	0.09**	0.09***
	(0.04)		(0.04)	(0.04)	(0.03)
Credit (0–3)	−0.00		−0.04		
	(0.06)		(0.06)		
Management score (0–29)	0.01		0.01		
	(0.01)		(0.01)		
Observations	229	232	229	229	229
Adj R-squared	0.250	0.255	0.264	0.267	0.264
R-squared	0.361	0.337	0.377	0.371	0.360

Notes: Robust standard errors in parentheses clustered at the firm level, $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

The dependent variable is the sum of the standardized values of the 4 growth measures. The dependent variables are levels in the follow-up surveys and all regressions control for the baseline value of the dependent variable. The regressions in columns 1 through 4 include controls for owner age and female, manufacturer, retail, firm <5 years old, location in Kumasi, panel and wave 1 fixed effects. The regression in column 5 includes only location, sector, wave 1 and panel fixed effects.

Source: Authors' calculations based on survey data.

15 Because half of those completing the competition receive customized consulting and training, then the regressions on table 3 will raise the average growth rates of those presenting before the panel if training positively affects growth. We show below that training has no effect on firm growth.

Column 3 includes both the panel score and the three ability measures. The magnitude of each coefficient is reduced when both sets of variables are included together.[16] We explain an additional 1.6 percentage points of the variance in growth (1.1 percent accounting for the additional regressor), compared with the regression including only the survey measures. To simplify the comparison of the relative contribution of the panel and the survey, in column 4 we include only the ability measure based on education, Raven, numeracy and financial literacy scores.[17] Both the single survey ability measure and the panel score remain significant. The coefficient on the panel score is about 20 percent larger than the coefficient on the survey measure, but the standard deviation of the survey measure is 1.3 times as big as the panel score. Hence, a one standard deviation difference in the survey measure or the panel score shows an impact of similar magnitude on growth. The main conclusion from this analysis is that the panel score adds to the prediction of growth above and beyond what is picked up by the survey.

Column 5 of table 4 repeats the regression from column 3 using only the panel FE and sector dummies as controls. We see that the exclusion of several controls makes little difference to the results of interest. In sum, while the surveys explain more of the variance in subsequent growth than the panel scores, the panel scores add predictive power above and beyond the information collected in the survey.

Finally, we can ask whether the panels and surveys are better predictors at the top or the bottom of the distribution. That is, do panels (surveys) do a better job of differentiating within the bottom of the distribution, or within the top of the distribution? The modest sample size limits what we can say on this to some extent. But in regressions not shown on the table, we find that when we leave out the top quartile of the distribution by panel scores, both the panel scores and the ability measure significantly predict growth. However, when we leave out the bottom quartile of the distribution by panel scores, only the survey ability measure significantly predicts growth.

This suggests that panels are better at cleaving off the bottom of the distribution, but not so effective at distinguishing within the top group. There are two points to keep in mind with regard to this conclusion. First, it may be that if growth were measured over a longer time period, the panelists would prove effective at sorting out the upper part of the distribution as well. Second, the exercise is biased in favor of the survey measure, because the sample is truncated on the basis of the panel measure. When we conduct a similar exercise excluding the top and bottom quartiles of the distribution by the ability measure, we find that the panel measure significantly predicts growth in either subsample, while the ability measure significantly predicts growth only in the upper three quartiles. Nevertheless, we interpret this evidence as suggesting that surveys are better able to separate the contenders at the top of the distribution, and panels at the bottom.

IV. Training and Growth

As an incentive to participate in the business plan competition, we announced that some participants would be awarded a scholarship to an enterprise training course. The probability of receiving a scholarship is increasing in the ranking of the entrepreneur by the panel. Those in the top quartile of the rankings had a 75 percent probability of receiving training; those in the lowest quartile had only a 25 percent chance of receiving training; and those in the middle two quartiles had a 50 percent chance of receiving training. In all, 70 of the 140 firms completing the competition received training.

16 The panel score is positively correlated with all three of the survey measures, but not very strongly. The correlations range from 0.20 (intelligence) to 0.27 (credit).

17 A linear combination of the three ability measures (standardized) produces qualitatively similar results. On the other hand, years of schooling—the most widely available measure related to ability—is not significant when used in lieu of the broader ability measure. Owner age, sector, or location are not associated with growth either.

The training contained two parts. The first was a five-day course modeled on the International Labor Organization's *Improve Your Business* program and was offered by the most experienced provider of this program in Ghana. The second part of the program was more intensive and customized. We hired a local consulting firm with extensive experience working with the SME sector in Ghana. Consultants from the firm first conducted an individualized assessment of the needs of each of the 70 firms selected for the training. Based on this assessment, the consulting firm designed 18 modules, usually of one day's or two days' duration. Around 40 percent of the firms (27 of the 70) were provided with additional individualized consulting services to help them set up financial record-keeping systems appropriate for their specific circumstances. Among the 18 modules offered as a part of the specialized training, nine were delivered in Accra and 15 in Kumasi. The number of participants per course ranged from two to 12, with an average of seven. The modules offered and the number of participants in each city is shown in appendix S4.

Our goal was to provide 2.5 days of customized consulting services to the selected firms, but take-up was lower than we had anticipated. Sixty percent of those eligible for training (43 of 70) attended at least one module. Those attending at least one module attended just over three on average, though several attended five or more, and one attended 13 different modules. Among the treatment group, those rated more highly by the panel were more likely to participate in the training. Other covariates, including pervious participation in business training, are not significantly associated with attendance. The consultant also offered follow-up visits to the individual businesses, with 27 firms participating in these sessions.

Attendance rates are comparable to those found in other studies of training programs for micro and small enterprises. (See the training projects reviewed in McKenzie and Woodruff 2014.) The training was more intensive and more customized than typical microenterprise training programs. The second phase of the training ended in June 2011, just about a month before the first follow-up survey began. We do not expect the training to have an effect on enterprise growth so soon. We therefore focus on the second follow-up survey, which took place a year later. We face issues of statistical power; however, since the sample size is modest, and smaller than the initial design.[18]

Those rated more highly by the panels are more likely to be allocated training. Within strata, however, assignment to training is random. Hence in all regressions we control for strata fixed effects. We also control for ability, access to credit, and management practices scores. These three measures are collectively correlated with both the panel scores—and hence the strata—and with future growth.

The results from this exercise are shown in table 5. We use the aggregate measure of growth as dependent variable. The first column shows that, on average, training has a negative effect on firm growth, albeit significant only at the .15 level. We also note that training makes firm exit somewhat more likely. There are 14 firms in the first follow-up and 16 in the second that report no revenue in either of the two months prior to the survey. If we equate zero revenue with exit, we find that those assigned to training are 7 percentage points more likely to exit. Based on this, we can reject the idea that the predictive power of panel rankings on firm growth is due to correlation between panel rankings and treatment.

Next we ask whether training has a heterogeneous effect depending on the potential for growth. There are two reasons to be interested in this interaction. First, panel judges knew that the probability of receiving training was increasing in panel ranking. It follows that panel members may have ranked applicants on the basis of who they thought would benefit most from training, rather than of who they thought would grow the fastest. If true, this would be a concern for our interpretation of the panel ranking. Second, it is not clear where in the distribution of ability we expect training to have the largest effect. Perhaps those at the top already know what is being taught in the courses. Alternatively, those at

18 Offsetting this is the fact that the training was more intensive than initially planned.

Table 5. Training and Growth

	Aggregate Growth (1)	Aggregate Growth (2)	Aggregate Growth (3)	Aggregate Growth (4)
Training	−0.22	−0.21	−0.23	−0.12
	(0.15)	(0.15)	(0.14)	(0.11)
Training * panel score		0.07		0.04
		(0.15)		(0.11)
Training * Ability			0.06	
			(0.07)	
Top quartile of panel scores	0.39	0.38	0.43	0.19
	(0.36)	(0.36)	(0.36)	(0.23)
Middle two quartiles of panel scores	0.33	0.35	0.36*	0.24*
	(0.21)	(0.22)	(0.21)	(0.14)
Observations	112	112	112	229
R-squared	0.448	0.450	0.415	0.392

Notes: Robust standard errors in parentheses clustered at the firm level, $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

The sample in columns 1 and 2 is the second follow-up survey; in column 3, we use both the first and second follow-up survey. The dependent variable is the sum of the standardized values of the 4 growth measures. The dependent variables are levels in the follow-up surveys and all regressions control for the baseline value of the dependent variable. All regressions include conrols for owner age and female, manufacturer, retail, firm <5 years old, location in Kumasi, panel and wave 1 fixed effects.

Source: Authors' calculations based on survey data.

the bottom may not be sophisticated enough to absorb what is being taught. It is unclear a priori who would benefit the most from training.

The results (column 2) show a positive but insignificant interaction between training and panel score. The 95% confidence bounds are wide (−0.51, 0.07). Thus there is only a very weak suggestion of a more positive (or less negative) training effect on those ranked more highly by the panel. Column 3 allows the training effect to vary with the aggregate ability measure from survey data. We find a very similar result: training has a measured negative but insignificant effect on those with median ability, and the effects of training are increasing with ability, though imprecisely and insignificantly so. Column 4 then repeats the regression from column 2 using data from both of the follow-up surveys. We find a weaker effect of training—both in levels and in the interaction. This suggests that the effect of the training was small at the first follow-up, and that the movement toward negative effects strengthened over time. In results not shown on the table, we find no interaction between previous experience with training and the effect of training we provided as a part of the project.

V. Discussion and Conclusions

A large share of the labor force in low- and middle-income countries is self-employed. Identifying which of the vast number of micro entrepreneurs has the potential to expand has been viewed as something of a holy grail for researchers. We approach the sorting of the microenterprises in two ways. First, we consider which of a set of measures constructed from survey questions are associated with subsequent growth of enterprises. Second, we used panels of successful small business owners and consultants, organized through a business plan competition, to select firms that they expect will grow more quickly.

The psychology literature (e.g., Kahneman and Klein 2009) suggests that in circumstances where "experts" do not receive regular feedback about their choices, expert opinion may not be reliable. Most of the panelists were not in the habit of identifying small firms with a high growth potential, and probably did not receive regular feedback on any such forecast they might have made in the past. In spite of this, we find that the experts' opinion has predictive power, even after we control for what we can learn

from surveys. One reason why experts' rankings are informative in this case may be due to the wide heterogeneity of growth potential among the population of small enterprise owners in low- and middle-income countries.

As an incentive to participate, applicants presenting before the panel were given a chance to win customized consulting for their business. By randomizing the training treatment across all firms, the experiment was designed in such a way as to be able to identify firms that benefit most from training. Unfortunately, the training we offered—which we endeavored to be the best training that could be offered in Ghana at the time—turned out to have little effect on firm performance, making ex post identification of predictors of high-marginal-return candidates impossible. What remains unclear is whether treatment has no measureable effect on subsequent firm performance because the vocational training was not useful or because it was useful but untreated top performers acquired equivalent vocational skills elsewhere on their own. In the latter case, treatment could still be beneficial because it speeds up acquiring useful skills, or reduces the cost of acquiring them—but this kind of benefit is difficult to capture with endline retrospective questions and was not measured in our experiment. As a result, in spite of our initial intentions, we are not able to say whether training should optimally be targeted at top performers or weaker performers. But the results do provide some guidance on selecting firms with the potential for faster growth.

References

Cabral, L., and J. Mata. 2003. "On the Evolution of the Firm Size Distribution: Facts and Theory." *American Economic Review* 93 (4): 1075–90.

Davis, S. J., J. Haltiwanger, R. S. Jarmin, C. J. Krizan, J. Miranda, A. Nucci, and K. Sandusky. 2007. "Measuring the Dynamics of Young and Small Businesses: Integrating the Employer and Nonemployer Universes." NBER Working Paper w13226. Cambridge, MA: NBER.

De Mel, S., D. McKenzie, and C. Woodruff. 2010. "Who Are the Microenterprise Owners? Evidence from Sri Lanka on Tokman v. de Soto," in *International Differences in Entrepreneurship*, Lerner, and Schoar, eds., Chicago: University of Chicago Press.

———. 2012. "Business Training and Female Enterprise Start-up, Growth, and Dynamics: Experimental Evidence from Sri Lanka." Mimeo. World Bank.

Fafchamps, M., and S. Quinn. 2015. "Aspire." NBER Discussion Paper 21084. Cambridge, MA: NBER.

Kahneman, D., and G. Klein. 2009. "Conditions for Intuitive Expertise: A Failure to Disagree." *American Psychologist* 64 (6): 515–26.

Klinger, B., and M. Schündeln. 2011. "Can Entrepreneurial Activity Be Taught? Quasi-Experimental Evidence from Central America." *World Development* 39 (9): 1592–610.

McKenzie, D. 2015. "Identifying and Spurring High-Growth Entrepreneurship: Experimental Evidence from a Business Plan Competition." World Bank Policy Research Working Paper 7391. Washington, DC: World Bank.

McKenzie, D., and C. Woodruff. 2014. "What Are We Learning from Business Training and Entrepreneurship Evaluations around the Developing World?" *The World Bank Research Observer* 29 (1): 48–82.

Shanteau, J. 1992. "Competence in Experts: The Role of Task Characteristics." *Organizational Behavior and Human Decision Processes* 53: 252–62.

Sutton, J., and B. Kpentey. 2012. *An Enterprise Map of Ghana*. London: International Growth Centre.