American Economic Review 2022, 112(3): 861–898
https://doi.org/10.1257/aer.20200751

Targeting High Ability Entrepreneurs Using Community Information: Mechanism Design in the Field†

By Reshmaan Hussam, Natalia Rigol, and Benjamin N. Roth*

Identifying high-growth microentrepreneurs in low-income countries remains a challenge due to a scarcity of verifiable information. With a cash grant experiment in India we demonstrate that community knowledge can help target high-growth microentrepreneurs; while the average marginal return to capital in our sample is 9.4 percent per month, microentrepreneurs reported in the top third of the community are estimated to have marginal returns to capital between 24 percent and 30 percent per month. Further we find evidence that community members distort their predictions when they can influence the distribution of resources. Finally, we demonstrate that simple mechanisms can realign incentives for truthful reporting. (*JEL* D82, G21, I38, L25, L26, O12, O16)

Not everyone has what it takes to be a successful entrepreneur. Numerous experimental studies of microentrepreneurs in the developing world find widely heterogeneous returns to cash and credit.[1] Yet governments, lenders, and nongovernmental organizations often lack hard information with which to target resources to high-growth entrepreneurs. This may be an especially pressing need given the scale of cash transfers—both through grants and loans—distributed to the poor in the developing world. For instance, a recent World Bank estimate suggests that over 700 million people in the developing world receive some kind of cash transfer from

* Hussam: Harvard Business School (email: rhussam@hbs.edu); Rigol: Harvard Business School (email: nrigol@hbs.edu); Roth: Harvard Business School (email: broth@hbs.edu). Stefano DellaVigna was the coeditor for this article. We would like to first thank our team for their tireless work on this project, especially our research manager Sitaram Mukherjee and our project assistants Prasenjit Samanta and Sayan Bhattacharjee. We are very grateful to Rohan Parakh for exceptional advice, research assistance, and field management. We are also grateful to Namita Tiwari, Suraj Jacob, and Meghana Mugikar for excellent assistance in the field. We thank Savannah Noray for meticulous technical assistance in Cambridge. We thank Shreya Chandra and Ana Paula Franco for help in preparing the paper for publication. This research was made possible with funding from the Asian Development Bank, Weiss Family Fund, PEDL, and the Schultz Fund at MIT. We received valuable feedback about this project from Abhijit Banerjee, Arielle Bernhardt, Pascaline Dupas, Peter Hull, Dean Karlan, David McKenzie, Benjamin Olken, Rohini Pande, Jonathan Roth, Simone Schaner, Christopher Woodruff. This project is registered in the AEA RCT Registry (AEARCTR-00011095); see Hussam, Rigol, and Roth (2016) for further details. This study received IRB approval from MIT under the COUHES Protocol 1403006218.

† Go to https://doi.org/10.1257/aer.20200751 to visit the article page for additional materials and author disclosure statements.

[1] For example, Fafchamps et al. (2014); de Mel, McKenzie, and Woodruff (2008); Banerjee, Karlan, and Zinman (2015).

their government.[2] Governments aiming to stimulate the local economy may desire to target cash transfers to high-growth microentrepreneurs.[3] Similarly, microfinance institutions have distributed loans to more than 139 million households in the developing world.[4] While microfinance institutions typically enjoy very high rates of repayment, there is increasing recognition that identifying high-growth entrepreneurs would allow them to offer more personalized forms of credit tailored to their borrowers' needs.[5] In this paper we argue that harnessing community information directly from a microentrepreneur's peers may provide a viable approach to identifying high-growth microentrepreneurs.

Our argument has three parts. First, we demonstrate that entrepreneurs in peri-urban Maharashtra have high quality information about one another along a variety of dimensions including marginal returns to capital. Their information is valuable for identifying high-growth microentrepreneurs even after controlling for a wide range of demographic and business characteristics. Second we demonstrate that entrepreneurs manipulate their reports to favor themselves, their friends, and their family when the distribution of resources is at stake. Finally we identify several simple techniques motivated by mechanism design that effectively realign incentives for accuracy.

Specifically, we conducted a field experiment with 1,345 entrepreneurs from Amravati, a city in Maharashtra, India. We assigned respondents and their nearest neighbors to peer groups of five people. After collecting detailed baseline data from all respondents, we asked entrepreneurs to rank their peer group members on predicted marginal returns to capital, profits, and other firm, owner, and household characteristics. Once the community reports were complete, we randomly assigned US$100 grants to one-third of entrepreneurs in order to induce business growth and assess the accuracy of respondents' predictions. We evaluate the accuracy of community information by comparing how well the rankings predict individuals' true outcomes as reported at baseline or in subsequent follow-up surveys.

Our first main finding is that community members can identify high-return entrepreneurs. While the average marginal return to the grant was about 9.4 percent per month, our point estimates of the marginal returns to capital of entrepreneurs ranked in the top third range from 24 percent to 30 percent. Had we distributed our grants using community reports instead of random assignment, we would have roughly tripled the total return on our investment.

To benchmark the value of community information, we compare its predictive accuracy against that of observable entrepreneur characteristics. We find that observable characteristics are indeed strong predictors of marginal return to capital.[6]

[2] Honorati, Gentilini, and Yemtsov (2015).

[3] The majority of government transfers are not explicitly earmarked for entrepreneurship. But as of 2020, the World Bank estimates that 81 percent of all employed people in low-income countries are self-employed (see https://data.worldbank.org/indicator/SL.EMP.SELF.ZS [accessed November 2020]). Therefore identifying who among them can put capital to productive use may be a high priority for determining the distribution of government transfers.

[4] Convergences (2019).

[5] See, e.g., Jayachandran (2020). Further, Rigol and Roth (2020) finds evidence that microfinance loan officers have valuable information for differentiating amongst which borrowers would benefit from larger more flexible loans.

[6] This stands in contrast to the findings of McKenzie and Sansone (2019), which documents another attempt to predict business growth using observable characteristics.

However, when we estimate marginal returns based on community information and control for a wide range of observables, we still find that those in the top tercile of the community prediction distribution earn 18–36 percentage points higher monthly returns than those in the bottom tercile, holding fixed other characteristics (for many configurations of observables, entrepreneurs predicted to be in the bottom tercile have negative marginal return to capital). This finding suggests that community information is valuable above and beyond information that can be captured by observables.

Our second main finding is that strategic misreporting is a first-order concern when eliciting community information. By random assignment, half of respondents were told that their reports would be used only for research purposes (the "no stakes" treatment) and the other half were told that their reports would be used to allocate US$100 grants to members of their community (the "high stakes" treatment). The correlation between community reports and true outcomes is on average 27 percent to 35 percent lower when allocation of resources is at stake, which significantly lowers the value of peer elicitation. We also identify who benefits from misreporting and by how much: we quantify the extent to which participants favor themselves, their family members, and their close friends (as identified by other group members).

Given the importance of strategic misreporting, we explore whether it is feasible to realign incentives to report truthfully.[7] Alongside the high stakes treatment, we cross-randomized treatments that varied respondents' immediate benefit (or cost) for truthful responses. Respondents were assigned to report either in private or in a public setting, with their fellow neighbors observing their reports. Participants were also randomly assigned to receive monetary payments based on the truthfulness of their reports. Payments were calculated using the *robust Bayesian truth serum* (RBTS), a peer prediction mechanism that determines participant scores as a function of the contemporaneous reports of other respondents (Witkowski and Parkes 2012).

Our third main finding is that methods grounded in mechanism design theory can be used to design a peer-elicitation environment in which truth telling is incentive compatible. Monetary payments and public reporting do little to improve the accuracy of self-reports. But payments substantially increase the predictive power of reports that entrepreneurs make about other group members. We provide direct evidence that monetary payments reduce the likelihood that respondents favor their family members or their close friends. Finally, we find that public reporting increases the predictive accuracy of reports about others when there are no stakes, but has no effect in a high stakes setting. This nuanced finding may reflect a heterogeneous treatment effect, or a noisily estimated impact of observability on the quality of reports.

Beyond targeting cash grants to high-growth microentrepreneurs, the methods in this paper may prove useful in other contexts. We provide an experimental framework for predicting heterogeneous treatment effects *before* treatment implementation. Namely, by asking subjects of the experiment to predict their own and their peers' treatment effects, researchers can leverage information embedded in their experimental contexts. This may serve as a complement to recently developed techniques to estimate heterogeneous treatment effects after the experiment is complete

[7] For theory on eliciting reports about community members on a network, see Bloch and Olckers (2019).

using observable characteristics (e.g., Wager and Athey 2018). Dal Bó et al. (2018) employs a similar experimental design.

Our findings contribute to several literatures. The idea that social networks—friends, family, colleagues—are a rich source of information has deep roots in development economics. Of particular relevance are the studies that use community reports to inform policy, broadly construed (e.g., Alatas et al. 2012; Beaman and Magruder 2012; Giné and Karlan 2014; Bryan, Karlan, and Zinman 2015; Basurto, Dupas, and Robinson 2019). Relative to this literature, our findings provide new insight into the depth and breadth of social knowledge contained in rural and peri-urban networks. We demonstrate that community members can predict marginal returns to capital, a metric that is difficult to estimate even using rich observables or expert opinions. This is evidence that community members have accurate knowledge of one another that is much deeper than what has been previously documented. In this sense our paper is related to Maitra et al. (2017), which leverages information from local traders to identify high return entrepreneurs; Beaman et al. (2020) and Barboni and Agarwal (2020), both of which examine the extent to which borrowers utilize information about their own marginal return to capital to inform their borrowing decisions; and Fafchamps and Woodruff (2017), which demonstrates that a panel of judges can identify high-growth entrepreneurs through a business plan competition.

We also contribute to a young literature that addresses strategic misreporting in targeting programs. Alatas et al. (2019) examines whether elite capture poses a problem for community reporting, but incentives to manipulate the distribution of resources may extend beyond community elites. Though Alatas et al. (2019) concludes that elite capture is not a significant concern, we find that misreporting is common when community members are told that their reports will influence distribution of grants. Importantly, we find that community members distort their reports in favor of their family and friends, rather than toward community elites. Within the microfinance context Rigol and Roth (2020) provides evidence that loan officers strategically withhold valuable endorsements for targeting larger more flexible loans to their borrowers unless they are provided with a compensation structure that rewards doing so.

Finally, our paper relates to the literature studying cash grants to microentrepreneurs across the developing world (e.g., de Mel, McKenzie, and Woodruff 2008; McKenzie and Woodruff 2008; Fafchamps et al. 2014). The emerging consensus is that the average marginal return to capital among microentrepreneurs is high. However, whether there are robust predictors that identify which entrepreneurs are likely to have high marginal return to capital and which not—or even whether there is meaningful variation in ex ante expected marginal return to capital—remains unresolved. Our paper demonstrates that microentrepreneurs have widely varying marginal return to capital ex ante, and that community information is a strong differentiator amongst microentrepreneurs based on their marginal return to capital.

I. Study Sample and Context

Our study takes place in Amravati, a city of about 550,000 people in the state of Maharashtra, India. Households in our sample come from nine neighborhoods along the perimeter of Amravati; we selected these neighborhoods because they have a relatively high proportion of microentrepreneurs.[8] These are densely packed peri-urban slums; in each of these neighborhoods, there are roughly 900 household dwellings in a 500 by 700 foot area. In September 2015, we conducted a complete door-to-door census of these neighborhoods, which encompassed 5,573 households. Based on households' responses to the census, we determined their eligibility for the study. In line with selection criteria of other recent "cash-drop" experiments (see, e.g., de Mel, McKenzie, and Woodruff 2008), all households in our sample have at least one enterprise with (i) US$1,000 or less in total working and durable capital and (ii) no paid, permanent employees.[9] Almost 30 percent of households in these neighborhoods owned at least one business and were eligible (1,576 households). Entrepreneurs in 1,345 of these households agreed to participate in our study so our sample population is reasonably representative of the universe of eligible enterprises in Amravati.

Characteristics of Microenterprise Owners.—Entrepreneurs in our sample are on average 41 years old and have roughly 7 years of formal education. Approximately 60 percent are male and almost all are married. Most entrepreneurs operate their business close to home, but they operate across a wide range of activities. About 33 percent of sample entrepreneurs work in manufacturing, typically as a tailor or stitcher. Another 31 percent work in services, mainly in food preparation and hair salons. A further 33 percent work in retail, most commonly running a grocery shop. Outside of these three sectors, entrepreneurs are spread evenly across construction and livestock rearing. On average, sample entrepreneurs earn profits of 4,983 rupees per month (about US$2.5 per day), which accounts for roughly half of their household income.

Characteristics of Microentrepreneurs' Peer Networks.—In order to elicit entrepreneurs' knowledge of one another, we assigned study participants to peer groups of five people based on geographical proximity. Peer groups are the unit of information collection: entrepreneurs are asked to report on only themselves and their other group members, rather than on the entire community. Importantly, we find that peers know their group members well. On average, peers reported that they visited another group member on 22 occasions in the previous 30 days. Respondents were unable to identify another group member in less than 1 percent of cases. Two-thirds of respondents identify at least one other group member as a family member or close friend. In 70 percent of groups, at least two people operate a business in the same (broad) industry category. Entrepreneurs also actively maintain strong social ties within their group: over 50 percent of respondents reported that they regularly discuss private family and business matters with at least one other group member.

[8] Our selection of neighborhoods was based on advice from local officials in the district collector's office. The nine neighborhoods are Belpura, Vilash Nagar, Mahajan Pura, Akoli, New Saturna, Old Saturna, Wadali, and Pathan Chawk.

[9] Following de Mel, McKenzie, and Woodruff's (2008) selection criteria, we excluded farmers and self-employed service people, such as domestic helpers and teachers. If there were multiple business owners in the household, we required that the household have at most US$2,000 in combined business capital.

II. Experimental Design

Below we explain experimental design and that data collected (Hussam, Rigol, and Roth 2016).

A. *Design of the Peer Elicitation Exercise*

Recruitment.—In October 2015, we visited the 1,576 eligible households and invited them to participate in our study. At the time of recruitment, households were told that a research team was conducting a project to study entrepreneurship and business growth.[10] In December 2015–April 2016, we conducted baseline surveys of the 1,345 sample households. Separately, we also assigned respondents to groups of five based on geographic proximity, for a total of 274 groups across all neighborhoods.[11] Once all baseline surveys in a given neighborhood were complete, surveyors returned to sample households to invite respondents to a meeting at the local town hall. Respondents were not given any information regarding the content of the meeting, or that they would be placed into groups with their peers. They were told, though, that to thank them for their participation in the study the research team would conduct a public lottery where some participants would be awarded a US$100 grant.

Explanation of the Exercise to Respondents.—Respondents who were assigned by the research team to the same group were asked to come to the town hall at the same time. Only one group was invited to the town hall at a time. Upon arrival at the town hall, respondents were each given 20 lottery tickets. They were told that, at the end of the activity, all people present would put their lottery tickets into an urn and grant winners would be selected by drawing lottery tickets. Participants were then separated and individually paired with a surveyor. Surveyors explained to participants that they would be asked to provide information about themselves and their neighbors. In order to ensure that participants were introduced to the elicitation exercise in a clear and consistent way, we created animated videos to introduce respondents to the concepts covered in the rankings questions and to guide them through the activity. When explaining the concept of marginal return to capital, we used examples to emphasize to respondents that an entrepreneur's projected marginal returns corresponds to their expected *change* in profits in response to the grant, and not their *level* of profits. After watching the videos, participants completed a series of quizzes to test their understanding of the activity and concepts. The introduction and subsequent ranking activity took place behind a privacy screen. The screen was there to ensure that coordination of responses would not be possible.

Questions Asked in the Ranking Exercise.—As a practice round, we first asked participants to rank themselves and their peers based on their level of education. We then asked respondents to rank themselves and their peers on predicted marginal

[10] No information regarding the community information nature of the project was disclosed to respondents at this time.

[11] We organized respondents into groups that would minimize the geographic distance between study households. The total number of respondents per neighborhood was not always a multiple of five, so some groups had four or six clients. Online Appendix Figure A1 shows the distribution of group size.

returns to a US$100 grant. Next we asked respondents to rank the group across several additional entrepreneur characteristics: average number of hours spent at work per week; performance in a digit span memory test; and, projected monthly profits six months postgrant disbursal, if the business owner were to receive a US$100 grant. We also asked about a number of household-level characteristics: average monthly income over the past year, total value of assets; total medical expenses in the past six months; loan repayment trouble over the previous year.[12] To minimize respondent fatigue, peer groups completed the ranking exercise only for a randomly assigned subset of these metrics (but all respondents completed the marginal returns ranking). For details on the sequence and selections of ranking questions by treatment group, see online Appendix D.

Zero-Sum Elicitation.—For marginal returns, business profits, and household income and assets, we asked all respondents to rank their peers using two methods: rankings relative to the particular members of the group, and reports placing each entrepreneur in quintiles relative to the community at large. The former has a zero-sum nature, in which promoting someone's position necessitates diminishing another's, and may therefore be more effective at inducing truthful reports (a respondent cannot merely place everyone in the highest position). However, if group members have correlated attributes, then these rankings may be less informative than rankings that assess each entrepreneur relative to the broader community. For all other questions we only collected relative rankings (this was also done to reduce fatigue).

Cross-Reporting.—In the spirit of cross-reporting techniques that play a prominent role in mechanism design and implementation theory (see Maskin 1999), we asked respondents to identify each group member's closest peer in the group, with the intention of exploring whether group members identified as close peers distort their reports to favor one another. We also ask respondents to identify who in their peer group has the most accurate information regarding each ranking metric. Implementation details are in online Appendix D.

B. *Description of Treatments*

Respondents were cross-randomized (at the group level) to give their ranking reports under the following three treatment conditions, forming a total of eight treatment cells: *NoStakes* versus *Stakes* (S_0 versus S_1), *Private* versus *Public* (P_0 versus P_1), and *NoPayments* versus *Payments* (T_0 versus T_1). The eight treatment conditions were randomized in clusters of eight groups. The clusters were based on geographic proximity. In the remainder of the paper we refer to these as randomization strata.

To economize on incentive payments, only the profits, income and assets questions were subjected to the stakes or incentive treatments; that is, some questions

[12] We hoped to have significant overlap in the sample with a local microfinance (MFI) institution and that we could use their administrative data to assess whether community members could predict the quality of borrowers. In the end, however, the areas in which we implemented the project did not have a significant presence of this MFI. So we drop this question from our analysis.

were always unincentivized and asked in a no stakes environment, regardless of treatment assignment. Online Appendix D details the sequence of questions and identifies which ones were subject to each treatment. When analyzing the effect of the three treatment conditions, regressions are always limited to the profits, income, and assets questions.

We also randomly selected one-third of our sample to receive US\$100 grants. Grant randomization occurred at the individual level and was stratified by group. See online Appendix Figure A2 for the randomization design.

High Stakes Environment (S_0 *versus* S_1).—All participants across treatment groups were given 20 lottery tickets upon arrival at the town hall. Respondents in the high stakes treatment were told that, for each question, the peer ranked highest (on average) by group members would receive extra lottery tickets, and so would have a better chance of winning the grant.[13] In order to ensure that we would have sufficient power to evaluate the quality of predictions from the marginal returns rankings, all participants completed this ranking in a no stakes setting (the marginal return ranking occurred prior to any mention of the high stakes treatment).[14]

Public Reporting (P_0 *versus* P_1).—Participants in both the public and private reporting groups responded to each ranking question behind a privacy screen, in the presence of only their surveyor. But in the public treatment, after completing each ranking question, peers came to the center of the room and sat in a circle with their response clipboard in front of them. Participants were told that they were doing this so that the survey coordinator could record their responses, but the primary purpose was to give them the opportunity to observe one another's rankings.[15] Crucially, participants understood ahead of doing the ranking exercise that their peers would see their responses. This was described to them in their introductory animation video and, to ensure that participants understood the setup, groups performed a practice round. In the privacy treatment, respondents completed all ranking questions before interacting with peers and, even after the activity was completed, group members did not see each other's individual responses.

Payments for Truthfulness (T_0 *versus* T_1).—The introductory video for participants in the monetary incentives group explained that they would be paid per ranking question, based on the truthfulness of their responses.

Payments were calculated using the RBTS (Witkowski and Parkes 2012). The RBTS requires eliciting not only respondents' first-order beliefs (i.e., what they believe is the answer to the question), but also their second-order beliefs (how

[13] We did not tell participants how many extra lottery tickets would be awarded to the person ranked highest; in order to keep the randomization as close to uniform as possible, we awarded only one extra lottery ticket per ranking. Respondents were in a high stakes setting for four ranking questions, and so a person in this treatment group could win at most four extra lottery tickets. Participants completed all rounds of ranking questions prior to the disbursal of the extra lottery tickets.

[14] Measures of profits among microentrepreneurs in settings like this one are notoriously noisy (see, for instance, de Mel, McKenzie, and Woodruff 2009). Due to budget constraints, our experiment is just powered to detect how well marginal returns rankings predict realized marginal returns when accuracy of reports is not confounded by the incentive to lie present in a high stakes setting.

[15] Surveyors report that respondents did in fact almost always look at their peers' rankings.

respondents believe others will answer the question). Importantly, under general conditions, discussed in Witkowski and Parkes (2012), and in our online Appendix C, RBTS provides incentives to tell the truth *without relying on ex post measures of accuracy*. This is a critical advantage in our setting, as we used RBTS to incentivize truthful reporting about marginal returns to capital but could not confirm these at the individual level.[16]

The principal drawback of RBTS is that its underlying logic is complex. As such we created an introductory video providing a basic overview of the payment rule and an explanation of the reporting requirements. We did not explain the details of the payment rule to participants.[17] Instead, participants were told that people who reported what they truly believed would receive an extra 100 rupees on average (which is equivalent to two-thirds of the average daily wage). Payments were calibrated using the empirical distribution of reports from Rigol and Roth (2017) to maximize the strength of the incentive to tell the truth while adhering to a project budget constraint. Rigol and Roth (2017) also provides evidence that respondents were likely to believe our assertion that telling the truth would maximize their expected payment.

Because we did not explain the details of the RBTS to respondents, we cannot be sure how their behavior would evolve as they learn about the payment rule from experience. Thus we view the variation in payments (T_1 versus T_0) as informative about whether accuracy is responsive to moderate monetary incentives for truthfulness, but not necessarily as informative about the optimal way to provide such monetary incentives. We discuss details of the RBTS scheme and its implementation in online Appendix C.

Groups that were not in the monetary payments treatment were given a lump sum payment to compensate them for their time.

Enterprise Grant.—Upon completion of the peer elicitation exercise, group members came to the center of the room and placed their lottery tickets into an urn. One respondent was blindfolded and then drew tickets to award US$100 grants to one or two group members (the number of winners per peer group was determined by random assignment, and households could win at most one grant). Prior to grant randomization participants filled out worksheets specifying how they would invest the grant if they won. Participants were encouraged to invest grant money into their enterprise although this was not enforced. Grant money was distributed to winners via bank transfer.[18]

Random assignment allows us to use the difference between postperiod profits of grant winners and postperiod profits of grant losers as an estimate of the average marginal return to the grant. We therefore identify the informational value of community members' reports by testing the predictive power of respondents' marginal return rankings against our estimates of true marginal returns. The details of our identification strategy are presented in Section IV.

[16] Our experimental design allows inference over average marginal return to capital for well-defined populations, but not for individual entrepreneurs.

[17] Indeed, Danz, Vesterlund, and Wilson (2020) provides evidence that respondents may provide more accurate reports when the details of a payment rule are obscured.

[18] All households in our sample had at least one member with a bank account prior to our study.

III. Data and Background Results

Description of the Data.—Our main analysis uses data from respondents' peer rankings during the elicitation exercise and from respondent surveys. Baseline surveys were conducted between December 2015 and April 2016, and four follow-up surveys were conducted between May 2016 and March 2017. After baseline and the elicitation exercise but before the first round of follow-up surveys, 8 households dropped out of the study (none of the households were grant winners). The remaining 1,337 households answered all follow-up surveys, save for one that did not answer our fourth follow-up, and one that did not answer our fifth follow-up.[19]

For all survey rounds, each business owner in the household completed a detailed business module about her own enterprise and answered questions about her well-being. The business module included questions on enterprise costs; revenues; profits; seasonality; inventories; labor inputs; assets; and business history.[20] At baseline, entrepreneurs also completed a digit span memory test and a set of psychometric questions.[21] In each survey round, the study respondent also provided information regarding her household's finances. The household-level module included questions on income, health expenditures, credit history and loan repayment issues, and assets. For the asset section, the respondent indicated whether the household owned a particular type of asset and its current resale value. Surveyors were trained to visually verify that the household owned each of the assets about which they reported. At baseline, the respondent also completed a full household roster with education and labor history for each household member. For a complete timeline of the project and data, see online Appendix Figure A3.

We note that as many of our key outcome variables are self-reported, they are measured with error. However, the random assignment of our treatments should assuage most concerns that measurement error will bias our experimental estimates. One potential concern of note is that respondents who overestimate their own profits on our surveys also systematically do so when they discuss business matters with their friends and neighbors (i.e., they are boastful). Consider a particular boastful respondent. If her peers predict that she has high profits, we may incorrectly conclude that they are well informed about her business, when in fact they are merely repeating the same biased estimates that she reported to our surveyors. This type

[19] One of the 1,337 entrepreneurs provided ranks only about themselves. So in specifications in which we omit the self-rank, the number of observations in the regressions reduce to 1,336.

[20] Entrepreneurs were instructed to net out wage payments to hired labor when calculating business profits, but the measure does not account for the value of the entrepreneur's own time.

[21] *Digit span memory test:* Surveyors conducted a memory test in which they showed the respondent a three digit number, put it away, and asked them to repeat the number back. They increased the number of digits until the respondent either could not recall the number correctly, or the respondent recalled ten digits correctly. The number of digits that a respondent was able to recall without error is their digit span score.

Psychometrics: Respondents answered each psychometric question in the module by providing their agreement with the given statement, where agreement was rated on a scale of one to five, with five indicating strong agreement and one indicating strong disagreement. A detailed description of the psychometric assessment module is in online Appendix D. The psychometric module questions are organized according to categories developed by industrial online psychologists: polychronicity measures the willingness to juggle multiple tasks at the same time (Bluedorn et al. 1999); impulsiveness is a measure of the speed at which a person makes decisions and savings attitudes (Barratt impulsiveness scale); tenacity measures a person's ability to overcome difficult circumstances (Baum and Locke 2004); achievement is a measure of satisfaction in accomplishing a task well (McClelland 1985); and locus of control measures a person's willingness to put themselves in situations outside of their control Rotter (1966).

of reporting error would introduce bias in our estimate of how much community members know about one another's income and profits, though not assets, which our surveyors were trained to independently verify. And, critically, if this type of measurement error is uncorrelated with their true profits (e.g., borrowers merely added a constant amount to their reported profits), it would not bias our estimates of community knowledge of marginal return to capital, as the random assignment of grants ensures that these boastful entrepreneurs will be equally represented in our treatment and control groups.

One important caveat is if boastfulness instead introduces an error term that is correlated with the receipt of the grant, then our estimates of the value of community information may be biased. Because the grant was randomly assigned, this correlation would need to be mediated by a correlation between true profits and the error introduced by boastfulness (e.g., the amount an entrepreneur inflates her reported profits scales with her true level of profits). In this case we could erroneously conclude that the community is well informed about marginal return to capital when in fact they are merely indexing on boastfulness. In other words, a group of boastful people, some of whom received a grant and some did not, will appear to have higher marginal return to capital than unboastful people, even if the true marginal return to capital is constant across all individuals. As we will show below, community information predicts marginal return to capital even controlling for a rich set of baseline characteristics including income and profits, and their interaction with the receipt of the grant. To the extent that entrepreneurial boastfulness is correlated with these characteristics, it is unlikely that our results are largely attributable to this type of reporting error. We formalize this discussion in online Appendix B.

Randomization Checks.—In online Appendix Table A1, we present the randomization check of baseline characteristics by treatment. To check for balance we estimate the model,

$$\text{(1)} \qquad Characteristic_{ij} \; = \; \tau_0 + \tau_1 Treatment_j + \gamma_r + \theta_m + \tau_s + \epsilon_{ij},$$

where i indexes the individual and j indexes the group; $Treatment_j$ is a dummy for whether the group was assigned to the *NoStakes* versus *Stakes* treatment (columns 1 and 2), the *NoPayments* versus *Payments* treatment (columns 3 and 4), the *Private* versus *Public* treatment (columns 5 and 6), and the *GrantWinner* versus *GrantLoser* treatment (columns 7 and 8); γ_r is the randomization stratum. We also add survey month (θ_m), and surveyor (τ_s) fixed effects. Standard errors are clustered at the group level.

The odd columns of online Appendix Table A1 show the mean of each characteristic for the control group in each block. So column 1 shows the means of characteristics for groups that were assigned to *NoStakes*. The even columns show τ_1 for each treatment (the difference between treatment and control characteristics). The characteristics in panel A are about the entrepreneur who was ranked during the ranking exercise and in panel B is the sector of her primary business. In panel C, we show household demographic characteristics as well as baseline average household income and the value of household assets. In panel D, we show household-level baseline business measures. All of the variables are aggregated over all household businesses. So if the ranked entrepreneur is the only business owner in the household, these reflect the values of only her businesses.

The majority of entrepreneur and household characteristics are balanced across treatment groups. At the bottom of the table, we present the F-test of whether the treatment group coefficients are jointly equal to zero. None of the joint tests of equality are rejected, suggesting that the randomization was effectively implemented.

A. *Community Knowledge about Households and Enterprises*

We begin our empirical analysis by investigating the depth of community members' knowledge of one another. As discussed in Section I, entrepreneurs have close social ties with peers in their neighborhood. In this section we show that community members have accurate knowledge about one another's concurrent household finances and enterprise characteristics. In our main empirical analysis (Section IV), we will argue that community members also make accurate forward-looking predictions about entrepreneurs' marginal returns.

During the ranking exercise, community members reported on their peers' average monthly household income, predicted monthly profits if they were to receive a US$100 grant, total value of household assets, household medical expenses over the previous six months, average weekly work hours, and predicted performance on a working memory test.[22] At baseline, we asked each entrepreneur to self-report answers to these same questions (at the time of the baseline survey, respondents had no knowledge of the purpose of the study or of the peer ranking activity). To evaluate the accuracy of community reports, we estimate the relationship between entrepreneurs' self reports and community members' reports for that person. We use the following regression model:

$$(2) \qquad Y_{ijq} \,=\, \beta_0 + \beta_1 \overline{Rank}_{ijq} + \gamma_r + \theta_m + \tau_s + \epsilon_{ijq},$$

where $\overline{Rank}_{ijq} \,=\, \sum_{k=1}^{n} \frac{1}{n} * Rank_{ikjq}$, n is the total number of group members in group j, and $Rank_{ijkq}$ is the rank that person k in group j assigns to person i (also in group j) on question q. So $\overline{Rank}_{ijq}$ is the average rank assigned to person i by the members of group j on question q. In our baseline specification $\overline{Rank}_{ijq}$ includes the reports of all community members, except for person i's report about herself. However for robustness we re-estimate all of our results using a variant of $\overline{Rank}_{ijq}$ that includes person i's rank about herself. These are included in online Appendix Table A2; they are qualitatively similar, though community predictions tend to be somewhat more informative when self-ranks are included.

The term Y_{ijq} is the corresponding outcome (baseline survey self report) for question q of person i; γ_r is the randomization stratum. To improve precision, we add survey month (θ_m), and surveyor (τ_s) fixed effects. Standard errors are clustered at the group level.

[22] We use a digit span test, which is a commonly used test for working memory. Respondents are shown flashcards with an increasing number of digits and asked to recall the numbers from memory. The surveyor records the total number of digits that the respondent correctly repeated back.

Table 1—What Respondents Know: Average Regressions in Levels

	Income (1)	Profits (2)	Assets (3)	Medical exp. (4)	Digit span (5)	Work hours (6)
Panel A. Average rank level						
Average rank	1,471.23	1,291.41	103,094.71	1,373.28	0.47	1.16
	(249.43)	(209.23)	(21,710.24)	(517.00)	(0.09)	(1.91)
Panel B. Average rank percentile						
Average rank	0.18	0.20	0.22	0.17	0.22	0.02
	(0.03)	(0.03)	(0.03)	(0.06)	(0.04)	(0.07)
Mean of outcome	8,833.84	6,913.14	475,362.21	2,866.78	5.19	61.32
	[6,846]	[6,011]	[719,309]	[5,389]	[2]	[23]
Observations	1,924	1,980	1,844	263	281	276
Number of households	1,029	1,039	997	263	281	276

Notes: Specification: This table estimates specification (2) in the paper. In panel A, average rank indicates the average ranking the entrepreneur was given by her peers for the question in the column heading. In panel B, average rank indicates the percentile of average rank level. The average rank is computed excluding a person's own self-rank. In columns 1, 2, and 3, the number of observations is greater than the number of households because we regress the outcome on both the zero sum (relative) and the nonzero sum (quintile) rank in a stacked regression and control for the ranking question. All respondents were asked to provide the quintile and relative rank for a randomly selected two of these three questions. A subset of respondents were also asked to provide the relative rank for the third question. A subset of respondents were also randomly selected to provide the relative rank for the questions in columns 4–6. Robust standard errors clustered at the group level in parentheses. All regressions include randomization strata, survey month, survey round, and surveyor fixed effects. The analog of this table that includes the self-rank can be found in online Appendix Table A2. Outcome variables: In panel A, the outcome variable is the level of the outcome labeled in the column header, as reported by the rankee at baseline. In panel B, the outcome variable is the percentile of the outcome in panel B. The number of observations varies across questions because each respondent answered only a subset of the questions as explained in Section IIA. For a description of the data that produced the outcome variables, see online Appendix D.

Table 1 presents the estimates of specification (2).[23] Panel A presents results in levels of the outcome and the average rank, so that a one unit increase in $\overline{Rank}_{ijq}$ is associated with a β_1 increase in the value of the outcome variable Y_{ijq}. To allow for comparability of estimates across questions, in panel B we convert each outcome and the corresponding average rank for each question into percentiles. Recall that while $Rank_{ikjq}$ is a discrete variable taking values from one to five, the *average rank* assigned to person i, $\overline{Rank}_{ijq}$ is a continuous variable. So, a 1 percentile increase in $\overline{Rank}_{ijq}$ is associated with a β_1 percentile increase in the outcome variable Y_{ijq}.

Entrepreneurs have substantial knowledge of their peers' household and enterprise characteristics. For example, in column 3 of panel B, a 1 percentile increase in the assets rank is associated with a 0.22 (standard error $=$ 0.03) percentile increase in the distribution of actual household assets. They can also accurately assess even difficult to observe characteristics: for instance, a one unit increase in the average rank is associated with a 0.47 (standard error $=$ 0.09) extra digits recalled in the digit span memory test (column 5 of panel A).

To contextualize the size of these estimates, we regress the business profits percentile on the percentile of the education of the entrepreneur and also the household assets percentile on the household income percentile: a 1 percentile increase in the

[23] In Table 1, we pool across all treatment groups: *NoStakes* versus *Stakes* treatment, the *NoPayments* versus *Payments* treatment, and the *Private* versus *Public*. In Sections IVE and IVF, we break these estimates up by treatment.

education distribution is associated with a 0.12 percentile increase in the distribution of business profits, and a 1 percentile increase in the income distribution is associated with a 0.33 percentile increase in the assets distribution.

IV. Main Results

A. *Entrepreneurs' Average Marginal Returns to Capital*

In this section we assess the average impact of the intervention on entrepreneurs' profits. Following de Mel, McKenzie, and Woodruff (2008), we estimate average marginal returns to the grant with the primary specification,

$$Y_{ijt} = \alpha_0 + \alpha_1 Winner_{it} + \phi_i + \delta_t + \theta_m + \tau_s + \epsilon_{ijt}, \tag{3}$$

where Y_{ijt} measures either total household business profits or household income of person i in survey round t.[24] We also present results limiting profits to the businesses owned by the entrepreneur ranked in the ranking exercise. We measure business profits by asking each entrepreneur in the household the following question: "Now that you have thought through your sales and your expenses from the past 30 days, I would like you to think about the profits of your business. By business profits, I mean taking the total income received from sales and subtracting all the cost of producing the items (raw material, wages to employees, fixed costs, etc). Can you tell me your business profits in the past 30 days?"[25] Household income is also measured using a single question: "What is your total household income over the past 30 days from all income generating activities?"

Like de Mel, McKenzie, and Woodruff (2008), we remove the outliers of the household income and total profits distributions (levels) by trimming the top 0.5 percent of both the absolute and percentage changes in profits measured from one period to the next. We also estimate regression specification (3) for $\log(Y_{ijt} + 1)$ of income and profits, using the untrimmed distributions.[26] In the main specification, we utilize three rounds of follow-up surveys, so t ranges from zero (baseline) to three.[27] We use $Winner_{it}$ as an indicator for whether person i won a grant at or before survey round t. We also include the following fixed effects: person (ϕ_i), survey round (δ_t), survey month (θ_m), and surveyor (τ_s). Standard errors are clustered at the group level. The coefficient of interest in regression specification (3) is α_1, which measures average marginal return to the grant in the sample.

[24] Bernhardt et al. (2019) reanalyzes data from several cash-drop experiments with microentrepreneurs and finds that measures of returns to capital differ substantially when analyzed at the household versus enterprise level. We therefore aggregate profits of all household businesses in our main specifications.

[25] De Mel, McKenzie, and Woodruff (2009) finds that asking one aggregate summary measure (rather than for the components) reduces noise in the estimation of profits.

[26] The results remain nearly identical whether we log transform the trimmed or untrimmed income and profits distributions. But because we use the log of the untrimmed distribution, the number of observations in the log specification is larger.

[27] We collected four rounds of follow-up data. We exclude the fourth round of data due to a financial shock that occurred right before our final round of data collection—the Indian government removed 1000 and 500 rupee bills from circulation, causing a major economic shock. We discuss this further in Section IVB and show that results are slightly noisier if we include this final survey round in the analysis.

In the *NoStakes* treatment group, assignment of grant winners was uniformly random: all participants received 20 lottery tickets and each group member was equally likely to have their tickets drawn from the urn. But, as described in Section IIB, respondents in the *Stakes* group were eligible to receive up to four extra lottery tickets, based on whether their peers ranked them highest for the treatment questions. To account for this, we weight all regressions by the inverse *propensity score*—i.e., the probability of being assigned to the relevant treatment (Rosenbaum 1987). In our setting, the probability of being assigned to treatment is fully determined by the number of lottery tickets that a subject receives, and the number of grants randomly allocated within each group. For instance, in a group with just one grant winner, the observation corresponding to respondent i who won the grant is weighted by i's inverse probability of winning the grant lottery, $\frac{\text{Total Tickets}}{\text{Tickets Held by Subject } i}$. And the observation corresponding to a respondent i who did not win a grant is weighted by i's inverse probability of losing the lottery, $\frac{\text{Total Tickets}}{\text{Total Tickets} - \text{Tickets Held by Subject } i}$. Observations from groups with two grant lotteries are handled similarly. This reweighting assures that the distribution of covariates is independent across treatment assignment (Austin 2011). In online Appendix Figure A4, we plot the distribution of lottery tickets in the sample.

Online Appendix Table A3 presents results from estimating specification (3). We find that the grant had a large positive effect on household income and total household profits. On average, households that win grants report an extra 566.5 rupees (standard error = 405.6) in household income and an extra 681.0 rupees (standard error = 319.0) in total household profits over households that were not awarded grants. These gains in household income and profits represent very high marginal returns to the grant: point estimates suggest that on average, households earn returns of 9.4–11.4 percent per month.[28] These estimates are in line with average returns estimated from cash grants in other settings: de Mel, McKenzie, and Woodruff (2008) finds marginal returns of 7.6 percent per month in response to a US$100 grant and Fafchamps et al. (2014) finds marginal returns of 9.7 percent per month in response to a US$120 grant.

B. *Can Communities Predict Entrepreneurs' Marginal Returns to Capital?*

In this section our measure of community knowledge is entrepreneurs' average marginal returns rank.

We collected these rankings both using relative rankings (how entrepreneurs compare to their peers within the group) and quintile rankings (how entrepreneurs compare to the broader community). In this section we utilize the quintile ranking responses, though the results are similar using the relative rankings. We discuss this further in Section IVB.

An entrepreneur's average marginal returns rank is the mean of all the ranks assigned to her by her group members. Since group members are in full agreement

[28] We arrive at this number by dividing the marginal increase in monthly income and profits by the size of the grant (6,000 rupees).

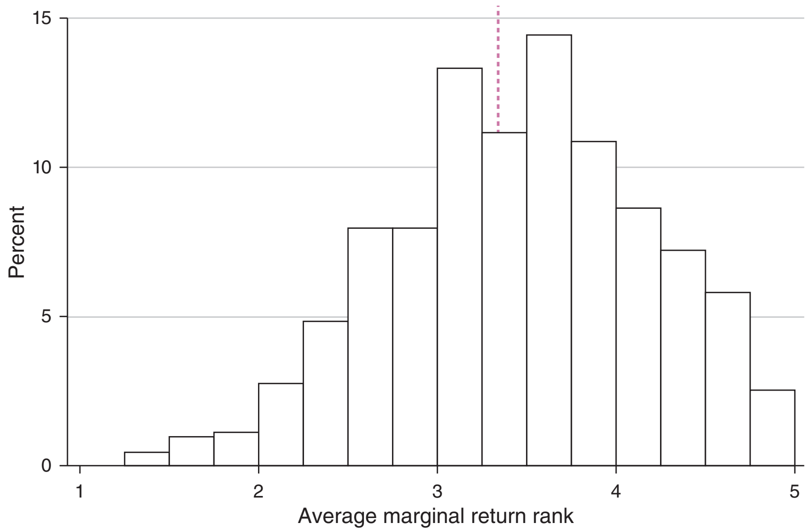

Figure 1. Distribution of the Average Marginal Returns Rank

Notes: This figure plots the distribution of the average marginal return quintile rank (nonzero sum rank). The average marginal return rank is the mean of every rank assigned to person *i* by all of her peers in her group.

about an entrepreneur's rank in fewer than 15 percent of cases, the distribution of average marginal return rank values is relatively smooth. We plot the distribution of average rank, which takes on values between one and five, in Figure 1.

In Figure 2, we plot kernel-weighted local polynomial regressions of log profits at follow-up by treatment assignment. The solid line plots a local polynomial regression of log profits on average marginal returns rank percentile (quintile ranks) for grant losers.[29] The dashed line plots a local polynomial regression of log profits on average marginal returns rank percentile (quintile ranks) for grant winners. The points in Figure 2 represent the average profits for grant winners and for grant losers at each two percentiles of the average marginal return ranks distribution.

We find that the distance between the two lines—in other words, entrepreneurs' marginal return to the grant—is increasing in the community's average ranking. An entrepreneur's marginal returns rank is strongly correlated with her increase in realized profits in response to the grant: for entrepreneurs in ranked in the bottom two-thirds of the ranks distribution, postgrant profits for winners and losers are statistically indistinguishable. But for entrepreneurs in the top third of the ranks distribution, the distance between treatment and control profits increases with marginal returns rank. In Figure 3, we replicate Figure 2 with baseline profits and show that differences in marginal returns to the grant are not driven by baseline differences in profits.

Figure 2 presents a joint confirmation of two hypotheses: that there is meaningful ex ante variation among entrepreneurs in terms of their expected marginal return to

[29] The average marginal returns rank percentile is the percentile of the average marginal returns rank distribution shown in Figure 1.

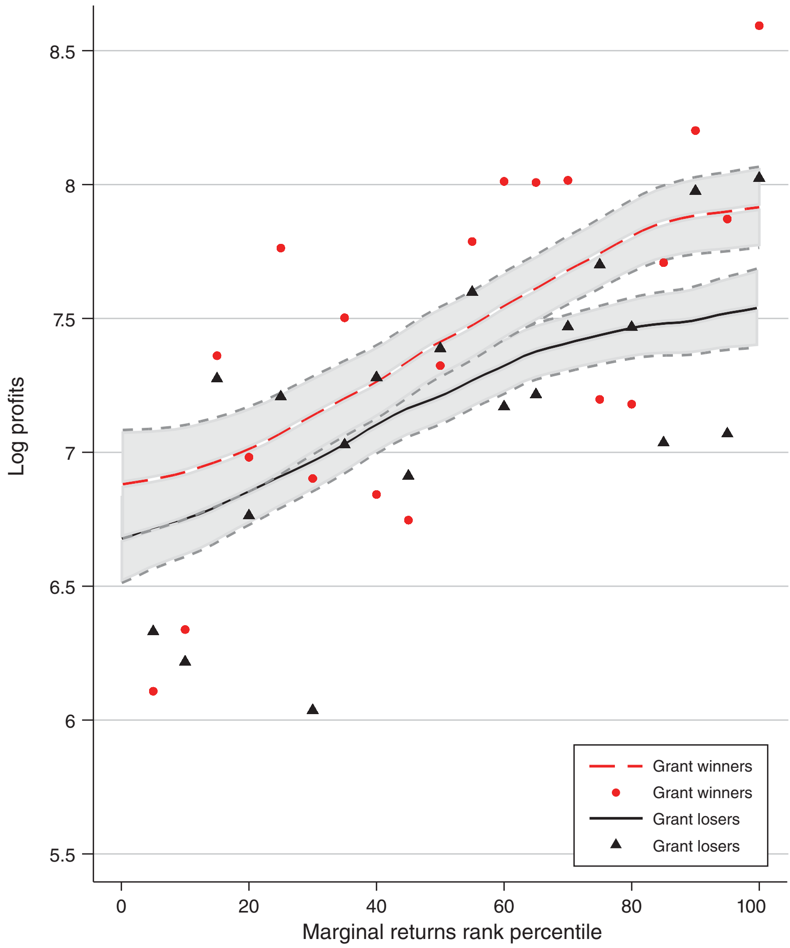

FIGURE 2. MARGINAL RETURNS TO THE GRANT BY PERCENTILE OF THE AVERAGE COMMUNITY RANKS DISTRIBUTION

Notes: This figure plots two kernel-weighted local polynomial regressions of log profits on the marginal returns rank percentile, estimated separately for respondents who won and respondents who did not win grants. Log profits is the average of the log value of profits in the post grant disbursal periods. The marginal returns rank percentile is the percentile of the average rank assigned to person *i* by all of her peers in her group (excluding the self-rank). Ninety percent confidence bands are shown. We additionally add a scatterplot of the data used to produce the local polynomial regression. Note that the scatterplot does not depict all of the data points used to produce the regressions. In order to make the figure readable, each point in the figure represents the average log profits for all of the entrepreneurs in the corresponding two marginal returns rank percentiles. So there is one point for every two marginal returns rank percentile for grant winners and grant losers.

capital, and that community members are able to accurately identify the ordering of their peers' heterogeneous returns ex ante. To quantify the size of these effects, we use a difference-in-difference specification and estimate of the relationship between community ranks and marginal returns to the grant. We extend the model from specification (3) to incorporate peer ranks:

$$(4) \quad Y_{ijt} = \alpha_0 + \alpha_1 \, Winner_{it} + \alpha_2 \, Winner_{it} \times \overline{Rank}_{ij} + \phi_i + \delta_t + \theta_m + \tau_s + \epsilon_{ijt},$$

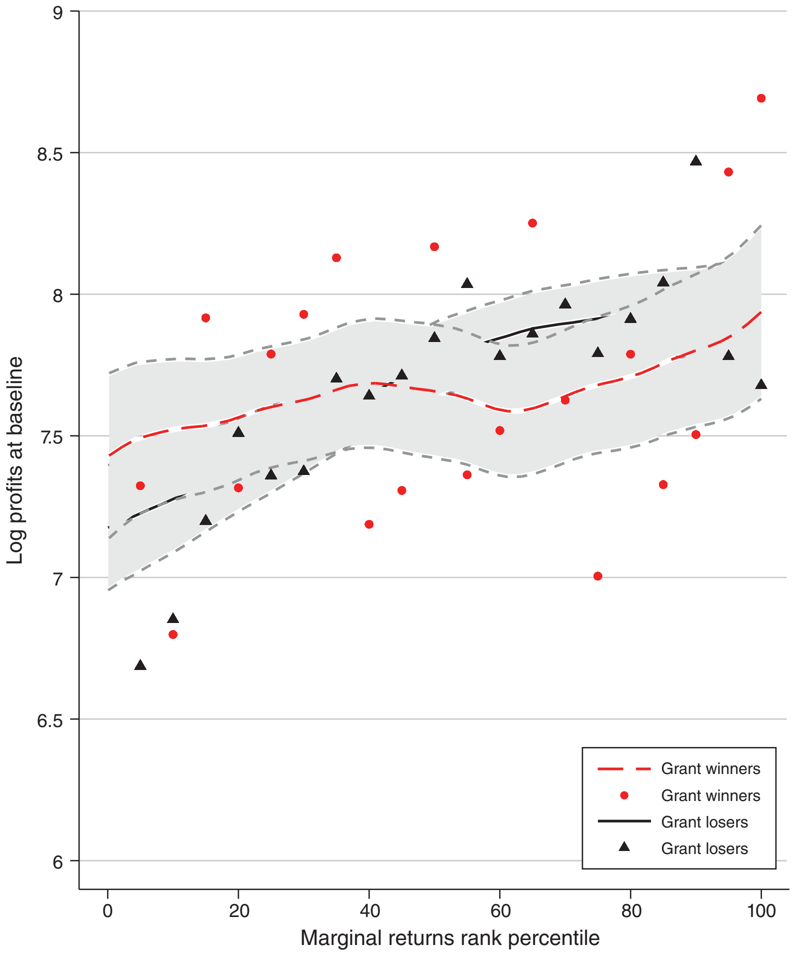

FIGURE 3. BASELINE LOG PROFITS BY AVERAGE COMMUNITY RANKS

Notes: This figure plots two kernel-weighted local polynomial regressions of log profits on the marginal returns rank percentile, estimated separately for respondents who won and respondents who did not win grants. Log profits is the log value of average profits at baseline. The marginal returns rank percentile is the percentile of the average rank assigned to person i by all of her peers in her group (excludes the rank). Ninety percent confidence bands are shown. We additionally add a scatterplot of the data used to produce the local polynomial regression. Note that the scatterplot does not depict all of the data points used to produce the regressions. In order to make the figure readable, each point in the figure represents the average log profits for all of the entrepreneurs in the corresponding two marginal returns rank percentiles. So there is one point for every two marginal returns rank percentile for grant winners and grant losers.

where $\overline{Rank}_{ij} = \sum_{k=1}^{n} \frac{1}{n} * Rank_{ikj}$, n is the total number of group members in group j, and $Rank_{ikj}$ is the rank that person k in group j assigns to person i (also in group j). So $\overline{Rank}_{ij}$ is the average marginal returns rank assigned to person i by the members of group j. The coefficient α_2 identifies the average additional marginal return to capital associated with a one unit increase in marginal return rank. The difference-in-difference specification estimates α_2 for a model in which marginal return increases linearly in the value of average rank. In our baseline specification we include all community reports in $\overline{Rank}_{ij}$, excluding the report of person i herself. We discuss how results are affected when we include an entrepreneur's report

TABLE 2—DO PEER REPORTS PREDICT TRUE MARGINAL RETURNS TO THE GRANT?

	Income (1)	Income (2)	Log income (3)	Log income (4)	Profits (5)	Profits (6)	Log profits (7)	Log profits (8)
Panel A. Average MR rank value								
Winner × rank	1,275.64	1,127.80	0.22	0.17	606.86	590.95	0.40	0.37
	(459.30)	(340.58)	(0.09)	(0.09)	(290.24)	(235.00)	(0.16)	(0.17)
Winner	−3,709.32		−0.62		−1,350.02		−1.04	
	(1,609.98)		(0.31)		(909.10)		(0.56)	
Panel B. Average MR rank tercile								
Winner × top tercile rank	2,261.13	2,167.94	0.34	0.19	1,301.83	1,107.33	0.67	0.48
	(802.98)	(627.62)	(0.21)	(0.19)	(557.19)	(404.91)	(0.31)	(0.31)
Winner × middle tercile rank	453.22	820.95	0.02	−0.00	118.19	139.78	0.07	−0.07
	(785.55)	(582.59)	(0.18)	(0.18)	(388.99)	(347.92)	(0.29)	(0.30)
Winner	−448.84		0.00		151.96		0.03	
	(622.35)		(0.16)		(374.89)		(0.25)	
p-value from F-Test								
Winner × top tercile rank = winner × middle tercile rank	0.026	0.034	0.062	0.245	0.027	0.029	0.023	0.039
Mean of outcome for grant losers	8,197.37	8,197.37	8.62	8.62	4,551.38	4,551.38	7.33	7.33
	[6,412.25]	[6412.25]	[1.35]	[1.35]	[5,159.23]	[5,159.23]	[2.55]	[2.55]
Controls		X		X		X		X
Observations	5,324	5,324	5,342	5,342	5,320	5,320	5,338	5338
Number of households	1,336	1,336	1,336	1,336	1,336	1,336	1,336	1,336

Notes: Specification: This table estimates specification (4) in the paper. Rank indicates the average ranking the entrepreneur was given by her peers for the marginal returns (MR) to grant quintile ranking (nonzero sum) question. It excludes the self-rank before producing the average ranking. See Figure 1 for a distribution of average rank. Top (middle) tercile rank is a dummy for whether the entrepreneur is in the top (middle) tercile of the average marginal return rank distribution. Winner indicates that the household is a grant recipient after baseline (after round 1 of data collection). The unit of observation is the household. Robust standard errors clustered at the group level in parentheses. All regressions include household, survey month, survey round, and surveyor fixed effects. The even columns also include all of the baseline controls in online Appendix Table A1 interacted with winner. All regressions are weighed by the inverse propensity score described in Section VA. Data in this table come from rounds 1 to 4 of data collection. Outcome variables: In columns 1 to 2 and 5 to 6 we show the *trimmed* distributions of income and profits, respectively, as described in Section IVA. In columns 3 to 4 and 7 to 8, we show the natural log of the (outcome + 1) of the *untrimmed* distribution (which is why the number of observations is greater than in the preceding column). For a description of the data that produced the outcome variables, see online Appendix D.

about herself in Section IVB. Motivated by the nonparametric estimates in Figure 2, we also estimate a nonlinear model in which the ranks distribution is divided into terciles and rank tercile is interacted (as above) with $Winner_{it}$. In online Appendix Table A4, we show that the sample is balanced across rank terciles and grant treatment groups at baseline.

Table 2 shows results of the difference-in-difference estimation of respondents' ability to predict true marginal returns to capital. Outcome variables are household income and total household profits, measured in levels and logs. Odd columns follow specification (4) while even columns also include a vector of interaction terms between $Winner_{it}$ and each of the control variables in online Appendix Table A1 (the uninteracted controls are subsumed in the person fixed effect ϕ_i). Therefore, even columns can be understood as measuring the value of community information in identifying high-growth entrepreneurs *over and above the predictive value of observable baseline characteristics.*

For the linear-in-rank version of the estimation (panel A), the coefficient α_2 is large, positive, and significant at least at the 10 percent level across seven of the eight columns. An extra unit of average rank is associated with increases in profits and income of between 591.0 rupees (standard error = 235.0) and 1,275.6 rupees (standard error = 459.3) per month, respectively. These amounts translate to increases in monthly returns to the grant of between 9.9 and 21.3 percentage points. Average marginal return to capital in the sample is between 9.4 and 11.4 percent per month and an entrepreneur ranked one standard deviation above the mean has monthly marginal return to capital between 18.7 and 24.7 percent (the mean and standard deviation of the marginal return rank are 3.34 and 0.73, respectively). For an entrepreneur ranked two standard deviations above the mean, monthly returns to capital are between 26.0 and 40.2 percent.

Panel B in Table 2 shows results from the nonlinear, tercile rank version of the difference-in-difference estimation. Consistent with results from the local polynomial regressions in Figure 2, we cannot reject that the entrepreneurs in the bottom tercile of the marginal returns rank distribution have zero returns to the grant.

Also consistent with Figure 2, while the coefficients for the middle tercile are typically somewhat larger than those for the bottom tercile, they are never statistically significant.[30] The strongest treatment effects of the grant are concentrated among entrepreneurs in the top tercile of the average rank distribution: depending on whether we use household income or profits, the coefficients on $Winner_{it} \times TopTercile_{ij}$ imply that monthly returns to the grant for the top tercile range from 24.2 to 30.2 percent. We can reject that the grant has the same effect for entrepreneurs in the middle and top tercile.

Comparing the odd columns to the even columns in panel A, we see that the estimates of α_2 are quite stable with respect to the inclusion of controls for observable characteristics, and their interaction with $Winner_{it}$. The same is true for the coefficients on $Winner_{it}$ interacted with tercile of rank in panel B. This suggests that community information is nonredundant with what can be inferred from observables regarding an entrepreneurs marginal return to capital. We further discuss the value of observables in identifying high-growth entrepreneurs in Section IVD.

These estimates indicate that the community has high quality information about which entrepreneurs can use a cash grant to grow their businesses. This could reflect knowledge that community members have about which entrepreneurs are the most talented. Or, this could reflect that entrepreneurs face heterogeneous credit constraints and that the community can identify which entrepreneurs are the most constrained. In either case, lending institutions or other organizations aiming to target capital to entrepreneurs with productive opportunities would have good reason to leverage community information.

Robustness Checks and Extensions

Evaluation of Community Information Using Cross-Sectional Variation: Regression specification (4) identifies the treatment effect of the grant off of the within-person differences in profits and income in the pre- and postgrant disbursal

[30] Mechanically, since the middle tercile is fixed, the difference between the level and log results occurs because there are some extreme right-tail observations in the distribution of income and profits for the middle tercile ranks. The weight of these outliers in the regression is diminished when the distributions are log transformed.

periods for grant winners and losers. As a robustness check, we also present results using an alternative specification in which the treatment effects are identified by comparing the cross-sectional differences between treatment and control groups in the postgrant disbursal periods, controlling for the baseline value of the outcome characteristic. The specification is

$$Y_{ijt} = \beta_0 + \beta_1 Winner_{ijc} + \beta_2 Winner_{ijc} \times \overline{Rank}_{ijc} + \beta_3 \bar{Y}_{ijPRE} + \gamma_r + \theta_m + \tau_s + \delta_t + \epsilon_{ijt}, \tag{5}$$

where Y_{ijt} is posttreatment outcomes (so t ranges from one to three rather than zero to three as in specification (4)) and $\bar{Y}_{ijPRE}$ is the pretreatment (time 0) value of the outcomes; γ_r is a randomization stratum fixed effect, θ_m is a survey month fixed effect, τ_s is a surveyor fixed effect, and δ_t is a survey round fixed effect. Standard errors are clustered at the group level. We present the analog of the average marginal returns (online Appendix Table A3) using specification (5) in online Appendix Table A5 and the analog of Table 2 in online Appendix Table A6.

Comparing Table 2 and online Appendix Table A6 we see that the point estimates in the linear specification (panel A) of online Appendix Table A6 are somewhat smaller than those in Table 2. In panel B, the coefficient on *top tercile rank* is somewhat smaller in online Appendix Table A6 than in Table 2 and the reverse is true for *middle tercile rank*. Nevertheless, our estimates using cross-sectional variation continue to suggest that community information is quite valuable. For instance, using our estimate of the coefficient on *top tercile rank* from column 1 of online Appendix Table A6, we see that community members predicted to be in the top third of the marginal returns distribution are estimated to have an average marginal return to capital of about 23 percent per month.

Utilizing Fewer Reports: Thus far in the analysis we have utilized the full set of peer reports in each group (excluding the self-rank). However, in settings where it is costly to collect additional reports it may be useful to gauge the marginal value of each incremental report (e.g., DellaVigna and Pope 2018). Online Appendix Table A7 reports specification (4) estimated for each of one, two, three, and four community reports.[31] The point estimate of the value of community information grows between 123 rupees and 280 rupees with each additional report, and each of estimates corresponding to the value of one, two, and three reports is statistically distinguishable from that of four reports. We find no evidence of diminishing returns to collecting additional reports up to four.

Self-Rank versus Community Ranks: Throughout the analysis so far, our measure of respondents' average rank excludes how they ranked themselves. The impact of including respondents' self-rank on community rank accuracy is ex ante ambiguous. We might expect entrepreneurs to have better knowledge about themselves

[31] We limit the analysis to reports made in groups of five respondents.

than they have about others. But respondents may also be more likely to strategically misreport in favor of themselves than when reporting about others.

Online Appendix Figure A5 replicates Figure 2 but uses the average marginal return rank percentile that includes the self-rank. In online Appendix Table A8, we replicate the results of Table 2 including the self-rank in our measure of average rank. Consistent with online Appendix Figure A5, we find that results are very similar to those presented in Table 2.[32] Comparing panel A of the two tables we see that adding self-rank to the linear specification slightly diminishes the point estimate of the value of community information, though not statistically significantly so. The estimates in panel B of the two tables are very similar, and in comparing them no clear pattern emerges. Therefore adding self-rank to the measure of community information does not seem to significantly improve the predictive power of community reports. This stands in contrast to the previous section, in which we established that adding additional community members to the average report did significantly improve its predictive power.

Overall, using estimates from column 1 in panel B, community members in the top third of the marginal returns distribution have an estimated 30.2 percent returns per month.

Quintile versus Relative Ranks: We collected both zero-sum and quintile community ranks; Section IIA contains a more detailed discussion of the two ranking methods. All analysis presented thus far in this section uses the (averaged) quintile community rankings. In online Appendix Table A9 we present the analog of Table 2 using the zero-sum rankings.

The community reports elicited using zero-sum rankings are somewhat less informative than those elicited using quintile rankings. For instance the point estimates in panel A of online Appendix Table A9 are smaller than those in Table 2. In panel B of online Appendix Table A9, the coefficient on *top tercile rank* is smaller that of Table 2 and the reverse is true for the coefficient on *middle tercile rank*. However the community report elicited through zero-sum rankings is still quite informative. For instance, drawing on the estimates in online Appendix Table A9, column 1, panel B, we see that those reported to be in the top tercile are estimated to enjoy an average marginal return to capital of 23.4 percent per month.

That community reports elicited using relative rankings are somewhat less informative than those elicited using quintile rankings might reflect that there are meaningful differences between entrepreneurs across groups (recall groups were selected so that members were geographically proximate to one another). Quintile rankings allow entrepreneurs to rank one another relative to the community at large rather than relative only to one another, and so may be preferred when attributes are correlated within members of the same group.

[32] To gauge the predictive value of self-rank judged against a single other randomly selected report about the same respondent, online Appendix Figure A6 replicates Figure 2, where profits are plotted against either self-rank or one randomly chosen group member's rank.

Other Measures of Community Information: Our primary measure of community information is the mean ranking assigned to a respondent by her group, but in principle there are many ways to summarize the information contained in community reports. Online Appendix Table A10 replicates Table 2 using the median community ranking rather than the mean. While the median community ranking is still predictive, it appears to be marginally weaker than the mean community ranking. Comparing the estimates in panel A of Table 2 and online Appendix Table A10, we see that those in the latter are smaller, and similarly for the coefficients on *top tercile rank*. Once again, however, the median community rank is still a strong predictor of marginal return to capital. Utilizing the estimate in online Appendix Table A10, column 1, panel B we see that those reported to be in the top tercile of median rankings enjoy an estimated average marginal return to capital of 22.8 percent per month.

Next we explore whether community reports with more agreement are more informative. Specifically, we augment specification (4) by including an interaction between the standard deviation of the community rank and its mean. The results, reported in online Appendix Table A11, do not provide strong evidence that the standard deviation of community ranking contains useful information over and above the mean, but the estimates are imprecisely estimated.

Individual versus Household-Level Profits—Following Bernhardt et al. (2019), our primary estimates utilize household income and profits, pooled across all sources within the household. In online Appendix Table A12, we present the results at the level of the client who was ranked by her peers. Point estimates and standard errors remain nearly identical, likely because 90 percent of households in our sample only operate one business.

Accounting for the Value of the Entrepreneur's Labor: Our measure of enterprise profits accounts for wages to paid employees but does not account for unpaid labor, and in particular does not account for the entrepreneurs' labor. In online Appendix Table A13 we present our main results for a measure of profits that attempts to account for the value of the entrepreneurs' labor. The details of how this measure is constructed are presented in online Appendix D. Comparing the estimates in online Appendix Table A13 to those in Table 2 we see that once we account for labor, the estimates remain very similar. Utilizing the estimate in column 1, panel B we see that those reported to be in the top tercile of rankings enjoy an estimated average marginal return to capital of 28.5 percent per month.

Demonetization: The month before we began our fifth (last) round of data collection, the Indian government removed from circulation two currency notes—the 1,000 and 500 rupee bills—overnight. The result was a tremendous shock to the formal and informal economy. As Chodorow-Reich et al. (2020) reports, traders experienced a 20 percent drop in sales due to demonetization. In fact, in the last round of surveying, over 50 percent of our sample reported being adversely affected by demonetization. For this reason, we exclude the postdemonetization wave of data from the analysis presented in the main tables. We replicate Table 2 with all five data rounds in online Appendix Table A14. The results are qualitatively similar but marginally noisier in a few specifications. Using the estimate in column 1 of panel B, those reported to be in the top tercile of rankings have an average marginal return to capital of 23.9 percent per month.

Value of Information in Groups of Five Members Only: The groups in our sample vary between four and six members, though 87 percent of the groups have five members (see online Appendix Figure A1). The interpretation of an entrepreneur's relative rank (though not quintile rank) depends on the size of her group. To ensure that none of our results are unduly influenced by including groups of varying size, in online Appendix Table A15 we re-estimate specification (4) restricting the sample to groups of only five members. The results are very similar, both qualitatively and quantitatively. For example, the average marginal return for the top tercile is 31.9 percent (column 1, panel B).

C. *Who Are the Top-Ranked Entrepreneurs and How Do They Invest Their Grant?*

In this section, we explore whether differences in entrepreneurs' characteristics and investment decisions can help explain the large gaps in returns that we observe.

Entrepreneurs' Investment Decisions in Response to the Grant.—In online Appendix Table A16 we limit the sample to grant winners and examine the relationship between self-reported grant investment decisions and marginal returns rank. Highly ranked entrepreneurs who won the grant report spending more of it on inventory and less of it on household expenditures, relative to lower ranked entrepreneurs who won the grant. Highly ranked entrepreneurs also appear to spend more of their grant on equipment, and to add more of their own money to business expenditures to complement the grant, though these differences are not statistically significant. While important differences in self-reported expenditures are evident, since money is fungible the observed effects might simply be due to mental accounting (see Karlan, Osman, and Zinman [2016] for evidence and implications).

To investigate whether grant investments translate to real increases in business inputs, we use regression specification (4) to compare inventories, business assets, and labor outcomes of grant winners and losers. Results are shown in Table 3. We find that the grant induces top and middle ranked entrepreneurs to accumulate higher capital stocks: relative to grant winners in the bottom tercile, top tercile grant winners report an extra 4,392.8 rupees (standard error = 2,688.9) worth of inventory and an extra 17,437.6 rupees (standard error = 8,107.2) of durable assets. The treatment increases the capital stock (inventory plus durable assets) of top tercile winners by approximately 199 percent of the grant amount. This treatment effect is within the confidence bound of increases in capital stock found in de Mel, McKenzie, and Woodruff (2008).

The grant also induces increases in an entrepreneurs' own labor, which may therefore be a complement to capital. In columns 3 and 4, we show that grant winners in the top tercile spend an extra 9.4 (standard error = 3.0) hours per week and an extra 4.3 (standard error = 1.3) days per month working when compared

TABLE 3—IMPACT OF GRANT ON BUSINESS INPUTS

	Business inventory	Durable business assets	Total hours worked past week	Total days worked past month	HH labor hours past week	HH labor wage bill past week	HH labor hours past week	Non-HH labor wage bill past week
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. Average MR rank value								
Winner × rank	1,100.669	11,105.691	5.548	1.715	2.535	11.174	3.108	148.370
	(1,747.836)	(6,586.042)	(1.600)	(0.642)	(1.261)	(6.575)	(2.236)	(113.841)
Winner	−2,305.989	−3.73e+04	−18.169	−4.442	−9.263	−36.601	−8.966	−334.730
	(5,151.955)	(21,307.082)	(5.590)	(2.228)	(4.221)	(34.447)	(7.000)	(392.939)
Panel B. Average MR rank tercile								
Winner × top tercile rank	4,392.837	17,437.565	9.406	4.303	6.036	23.884	6.087	189.237
	(2,688.881)	(8,107.212)	(3.028)	(1.336)	(2.759)	(12.198)	(3.629)	(178.265)
Winner × middle tercile rank	1,889.543	6,394.050	1.057	2.283	4.841	−18.042	0.347	67.758
	(1,299.603)	(8,676.901)	(3.159)	(1.308)	(2.310)	(35.042)	(3.623)	(247.019)
Winner	−955.397	−8,952.308	−3.490	−1.145	−4.808	−1.652	−0.985	65.456
	(1,317.737)	(5,652.467)	(2.554)	(1.074)	(2.120)	(13.895)	(2.424)	(163.909)
p-value from F-test Winner × top tercile rank = winner × middle tercile rank	0.313	0.348	0.006	0.087	0.568	0.276	0.196	0.655
Mean of outcome for grant losers	6,310.38	84,064.11	41.37	23.94	5.21	12.74	6.95	268.52
	[24,736.16]	[1,819,807.02]	[32.36]	[12.80]	[16.23]	[252.11]	[36.97]	[1,705.80]
Observations	5,284	5,319	5,254	5,254	2,649	2,649	2,649	2,649
Number of households	1,336	1,336	1,336	1,336	1,336	1,336	1,336	1,336

Notes: Specification: This table estimates specification (4) in the paper. Rank indicates the average ranking the entrepreneur was given by her peers for the marginal returns (MR) to grant quintile ranking (nonzero sum) question. It excludes the self-rank before producing the average ranking. See Figure 1 for a distribution of average rank. Top (middle) tercile rank is a dummy for whether the entrepreneur is in the top (middle) tercile of the average marginal return rank distribution. Winner indicates that the household is a grant recipient after baseline (after round 1 of data collection). The unit of observation is the household (HH). Robust standard errors clustered at the group level in parentheses. All regressions include household, survey month, survey round, and surveyor fixed effects. All regressions are weighed by the inverse propensity score described in Section VA. Data in this table come from rounds 1 to 4 of data collection. Outcome variables: The number of observations in columns 1–4 varies due to missing data across the rounds. Variables reported in columns 5–8 were only collected at baseline and in round 4. For a description of the data that produced the outcome variables, see the online Appendix D.

to winners in the bottom tercile.[33] Grant winners in the top tercile also utilize 6.0 (standard error = 2.8) more hours of household labor per week than winners in the bottom tercile.

Demographic Characteristics of Top-Ranked Entrepreneurs: In Table, A17 we compare baseline characteristics of households and entrepreneurs in all three terciles of the marginal returns ranks distribution. In column 1, we present the mean of each characteristic for the bottom tercile group. We then estimate the following model:

$$(6) \quad Y_{ijc} = \beta_0 + \beta_1 (MiddleTercile)_{ijc} + \beta_2 (TopTercile)_{ijc} + \gamma_r + \theta_m + \tau_s + \epsilon_{ijc}.$$

In columns 2 and 3, we present the coefficients from regressions of each baseline characteristic on whether the respondent is ranked in the middle (β_1) or top (β_2)

[33] See online Appendix Table A13 for estimates of specification (4) with a measure of profits that accounts for the entrepreneurs' labor.

terciles, respectively. Coefficients can be interpreted as the difference in each characteristic associated with being in one of the upper terciles relative to being in the bottom tercile.

When compared to bottom-ranked entrepreneurs, top-ranked entrepreneurs are 11 percentage points more likely to be male, have an extra 2 years of education, and are 1.6 years younger. They also remember an average of 0.59 digits more in the digit span memory test. We asked business owners how much a salaried job would have to pay per month in order for them to exit self-employment. Top ranked entrepreneurs report that they would require 23.3 percent higher monthly wages to leave their businesses. Households with a top-ranked entrepreneur look similar in terms of demographics, although top ranked households are slightly less likely to have a household member who is a daily wage worker. They have the same total number of businesses as households in the lower terciles. But these households have 36.0 percent more assets and their businesses earn 50.5 percent higher profits per month. They also earn 14.1 percent higher monthly income. For the most part, entrepreneurs in the middle tercile have baseline characteristic means that lie between the means of the bottom and top ranked entrepreneurs.

D. *Benchmarking the Value of Community Information against Observables*

We showed in the previous section that top-ranked entrepreneurs differ from low-ranked entrepreneurs across several observable demographic characteristics. These findings raise the question, are community members simply using observable information to rank one another? Our analysis from Section IVB suggests not. Specifically, recall that in Table 2, the estimates of the value of community information were stable with respect to the inclusion of a wide range of baseline demographic and business controls and their interaction with $Winner_{it}$. In online Appendix Table A18 we further control for psychometric characteristics and their interaction with $Winner_{it}$ (Klinger, Khwaja, and del Carpio 2013), and once again find that the estimates of the value of community information are stable to their inclusion.[34] This indicates that the information contained in the community reports is not largely overlapping with the information contained in baseline demographic and business characteristics with regards to predicting marginal return to capital. In this section we go one step further and compare the predictive power of each source of information.

To form a prediction of marginal return to capital based on observable characteristics, we estimate the following model:

$$Y_{ijt} \;=\; \alpha_0 + \alpha_1\, Winner_{it} + \alpha_2\, Winner_{it} \times X_i + \phi_i + \delta_t + \theta_m + \tau_s + \epsilon_{ijt}, \tag{7}$$

[34] Regressors are labeled according to the psychological trait for which they are meant to proxy (the specific wording of the statement is found in online Appendix D). There are two traits that are strongly predictive of marginal returns: optimism and achievement. We find that optimism negatively predicts marginal returns: business owners who are more likely to agree with the statements "In times of uncertainty I expect the best" and "I'm always optimistic about the future" and those who are more likely to disagree with "If something can go wrong with me, it will" have lower self-reported marginal returns. People who agree with the statement "Part of my enjoyment in doing things is improving my past performance" tend to have higher marginal returns.

where X_i is a vector of baseline characteristics for respondent i, and the rest of the variables are defined as above (note that the respondent fixed effect ϕ_i subsumes the uninteracted control vector X_i). To ensure the model is not overfit to the idiosyncratic features of our sample, the model is estimated using double-lasso (Belloni, Chernozhukov, and Hansen 2014), where the universe of potential characteristics included in X_i is all of those listed in online Appendix Table A1. We use the resulting model, combined with the vector X_i for each respondent, to form a prediction of each respondent's marginal return to capital based on observables. We denote each respondent i's predicted marginal return by $\hat{MR}_i^{Obs}$.

For comparability with our estimates of the value of community information, we then divide respondents into terciles based on their $\hat{MR}_i^{Obs}$. We then replicate the analysis presented panel B of Table 2, with our prediction based on observables. Specifically we estimate

$$(8) \qquad Y_{ijt} \;=\; \alpha_0 + \sum_{t=1}^{3} \alpha_t Winner_{it} \times Tercile_{ij}^{l} + \phi_i + \delta_t + \theta_m + \tau_s + \epsilon_{ijt},$$

where $Tercile_{ij}^{l}$ is a dummy indicating whether respondent i in group j falls into the lth tercile based on the prediction from observables and the rest of the variables are defined as above. The results are presented in the odd columns of Table 4.The controls selected for inclusion in the model are presented in online Appendix Figure A7. The point estimates indicate that observables are useful for predicting marginal return to capital, though the coefficient on *top tercile* is only statistically significant for the profits outcome variable. Comparing these estimates to those in panel B of Table 2 suggests that observables are about as informative as community rank; community rank appears to be a better predictor of income while observables perform better at predicting profits.

Next we investigate how much value community ranks add over and above observables in predicting marginal return to capital. Therefore we repeat the exercise above, now re-estimating specification (7), but where X_i includes not only the observable characteristics listed in online Appendix Table A1 but also community rank. We then form a prediction of each respondent's marginal return to capital based on observables and the community ranking, divide respondents into terciles and then reestimate the model in specification (8).

The result is presented in the even columns of Table 4. For all outcome variables, the prediction based on both observables and community information is stronger than the corresponding prediction based only on observables. With the exception of column 5 versus column 6, entrepreneurs that fall in the top tercile of the prediction based on both sources of information have statistically significantly higher marginal return to capital than those who fall in the top tercile of the prediction based on observables alone. For instance, looking at column 1, entrepreneurs who fall in the top tercile of the prediction based on observables alone enjoy a marginal return to capital of 13.6 percent per month. The corresponding estimate for entrepreneurs who fall in the top tercile of the prediction based on observables and community ranks is 38.2 percent per month (and from column 1 of panel B of Table 2 we see that the corresponding estimate based on community information alone is 30.2 percent per month). Therefore, even if a policymaker had access to the wide array of

Table 4—Observable versus Ranks Prediction

	Income (1)	Income (2)	Log income (3)	Log income (4)	Profits (5)	Profits (6)	Log profits (7)	Log profits (8)
Winner × top tercile controls	1,157.509		0.115		2,377.487		0.093	
	(752.152)		(0.202)		(608.675)		(0.311)	
Winner × top middle controls	1,576.349		0.206		1,599.643		−0.081	
	(868.320)		(0.200)		(498.874)		(0.276)	
Winner × top tercile controls + rank		3,559.464		0.632		2,752.701		0.798
		(725.716)		(0.180)		(569.789)		(0.302)
Winner × top middle controls + rank		1,867.939		0.326		1,288.719		0.247
		(792.343)		(0.164)		(423.688)		(0.246)
Winner	−342.438	−1,265.233	0.031	−0.180	−652.922	−656.104	0.309	−0.031
	(538.084)	(575.034)	(0.173)	(0.088)	(437.700)	(412.129)	(0.234)	(0.210)
p-value from F-test Winner × top tercile = winner × middle tercile	0.625	0.038	0.571	0.156	0.209	0.007	0.524	0.045
Mean of outcome for grant losers	8,197.37	8,197.37	8.62	8.62	4,551.38	4,551.38	7.33	7.33
	[6,412.25]	[6,412.25]	[1.35]	[1.35]	[5,159.23]	[5,159.23]	[2.55]	[2.55]
Observations	5,324	5,324	5,342	5,342	5,320	5,320	5,338	5,338
Number of households	1,336	1,336	1,336	1,336	1,336	1,336	1,336	1,336

Notes: Specification: This table estimates specification (8) in the paper. Top (middle) tercile controls is a dummy for whether the entrepreneur is in the top (middle) tercile of predicted marginal return to capital based on observables. Top (middle) tercile controls + rank is a dummy for whether the entrepreneur is in the top (middle) tercile of predicted marginal return to capital based on observables plus the average community ranking (excluding the entrepreneur's ranking of herself). Both predictive models were constructed using the process described in Section VD. Winner indicates that the household is a grant recipient after baseline (after round 1 of data collection). The unit of observation in the household. Robust standard errors clustered at the group level in parentheses. All regressions include household, survey month, survey round, and surveyor fixed effects. The even columns also include all of the baseline controls in online Appendix Table A1 interacted with Winner. All regressions are weighed by the inverse propensity score described in Section VA. Data in this table come from rounds 1 to 4 of data collection. Outcome variables: In columns 1 to 2 and 5 to 6 we show the trimmed distributions of income and profits, respectively, as described in Section VA. In columns 3 to 4 and 7 to 8, we show the natural log of the (outcome+ 1) of the untrimmed distribution (which is why the number of observations is greater than in the preceding column). For a description of the data that produced the outcome variables, see online Appendix D.

observable characteristics listed in online Appendix Table A1, community information would remain valuable.

E. *Do Peers Distort Their Responses When There Are Real Stakes?*

The analysis in the previous sections has shown that communities are well informed about members' marginal returns to capital. But to be of practical use, community members need to report their opinions *truthfully*.

In this section, we quantify whether and by how much community members distort their reports in high stakes settings.

The analyses in this section examine the relationship between community reports and entrepreneurs' business characteristics; income, assets, and profits. As explained in Section IIB, we did not randomize the *Stakes* and *NoStakes* treatments until after the marginal returns ranking was completed due to power considerations. As a result, we cannot include predictions about marginal return to capital in our analyses of incentives.

TABLE 5—DO RESPONDENTS DISTORT RESPONSES?

	Pooled questions (1)	Quintile questions (2)	Relative questions (3)	Pooled questions (4)	Quintile questions (5)	Relative questions (6)
Rank	0.164	0.150	0.177			
	(0.019)	(0.019)	(0.022)			
Rank × stakes	−0.056	−0.053	−0.057			
	(0.026)	(0.027)	(0.030)			
Average rank				0.237	0.230	0.245
				(0.026)	(0.030)	(0.028)
Average rank × stakes				−0.071	−0.078	−0.067
				(0.037)	(0.043)	(0.040)
Observations	22,526	10,514	12,012	5,748	2,685	3,063
Number of households	1,345	1,345	1,345	1,345	1,345	1,345

Notes: Specification: This table estimates specification (9) in the paper. The regressions include *Stakes*, but the coefficient is not reported in the table. In columns 1–3, rank is the percentile corresponding to the rank that person *i* in the group assigned to entrepreneur *j* in the group. So the unit of observation in these 3 columns is the ranker-rankee pair. Rank excludes the self-rank. In columns 4–6, average rank indicates the percentile of the average ranking the entrepreneur was given by her peers for a particular question. So the unit of observation is the rankee. Average rank excludes the self-rank. Robust standard errors clustered at the group level in parentheses. All regressions include ranking question, randomization strata, survey month, and surveyor fixed effects. The analog of this table that includes the self-rank can be found in online Appendix Table A19. Outcome variables: In columns 1 and 4, we pool across questions 1–3 in panel A of Table 1 (in order to be comparable across questions, the outcome variable is percentilized). In columns 2 and 5, we limit the analysis the quintile (nonzero-sum) questions. In columns 3 and 6, we limit the analysis to the relative (zero-sum) questions. So column 1 pools columns 2 and 3 together. Column 4 pools columns 5 and 6 together. The number of observations varies between columns 2 and 3 because each respondent answered only a subset of the questions as explained in Section IIA. For a description of the data that produced the outcome variables, see online Appendix D.

In order to assess whether and how peers lie when there is incentive to strategically misreport, half of our sample was informed that their rankings would affect the probability that their peers (or themselves) would win the US\$100 grant (this is the *Stakes* group). Respondents in the *NoStakes* group continued to believe that their ranking responses would only be used for research purposes. We assess strategic misreporting in Table 5 by amending specification (2) to compare accuracy in the *Stakes* and *NoStakes* groups:

$$Y_{ijq} = \alpha_0 + \alpha_1 \overline{Rank}_{ijq} + \alpha_2 Stakes_j + \alpha_3 Stakes_j \times \overline{Rank}_{ijq} + \gamma_r + \theta_m + \tau_s + \delta_q + \epsilon_{ijq}. \tag{9}$$

The model includes the following fixed effects: randomization stratum (γ_r), survey month (θ_m), and surveyor (τ_s). Standard errors are clustered at the group level. The variable α_1 captures the accuracy of the report in the control group (*NoStakes*); α_3 indicates the extent to which the rankings are differentially informative when respondents are told their reports will be used to help determine grant allocation.[35,36]

[35] To reduce clutter in the regression tables, we have omitted the *Stakes* coefficient from the regression report as it does not contain information relevant for the interpretation of results, but rather simply adjusts the constant.

[36] In this section, we pool across the *Public* and *Payments* treatments.

To increase power, we stack the percentilized outcomes and ranks across the first three columns presented in panel A of Table 1 and add a question fixed effect (δ_q) to the regression model.[37] The outcomes and rank variables are percentilized to allow for comparability across questions. Our measure of $\overline{Rank}_{ijq}$ excludes self-rank; the analog including self-rank is presented in online Appendix Table A19.

Respondents may have idiosyncratic preferences for misreporting about certain peers in their group and may otherwise make idiosyncratic errors. One way to reduce noise is to average across all reports given about a particular group member. So in columns 1–3 of Table 5, we show the regressions at the ranker-rankee level of observation $(Rank_{ijmq})$ and column 4–6 are the regressions with the average rank $(\overline{Rank}_{ijq})$. We observe that the average predictiveness of ranks in the $(NoStakes)$ group increases significantly when reports are averaged: in column 1, a 1 percentile increase in the rank distribution is associated with a 0.16 (standard error = 0.02) shift in the outcome distribution in the individual regressions and a 0.24 (standard error = 0.03) shift in the average regression (column 4).

Do respondents misreport in the high stakes settings? We find that the coefficient on $Rank \times Stakes$ is large, negative, and significant. We note that this was not ex ante clear: the *Stakes* treatment may have had a positive effect since introducing stakes may have caused respondents to focus or take the exercise more seriously. The regression implies that responses are significantly less accurate when respondents have an incentive to behave strategically: in the pooled individual regression in column 1, the responses become 34.1 percent less accurate in the *Stakes* group. The effect persists even when averaging reports; in column 4, the averaged responses become 30.0 percent less informative in the *Stakes* group.

Lastly, we asked respondents to rank their peers relative to others in the group (zero-sum ranking) and also relative to the community by reporting the quintile of the neighborhood distribution that they believe the peer to be in (quintile ranking). We hypothesized that quintile ranks could contain more valuable information about rankings because entrepreneurs are compared to the community more broadly than only the group. But they could also be more susceptible to misreporting: unlike with zero-sum ranks, respondents could, for example, place all of their peers in the top quintile of the distribution indicating that everyone is equally excellent.

To compare these two elicitation methods, in columns 2 to 3 and 5 to 6, we show the results by separately stacking zero-sum and quintile rankings. In all four columns, the outcome variable is the same (percentile of Y_{ijq}). What changes is the method of reporting. In columns 2 and 5, the regressor is the percentile in the (individual or average) quintile rank distribution. In columns 3 and 6, the regressor is the percentile in the (individual or average) zero-sum rank distribution.[38] The coefficients on *Rank* in the individual (columns 2 and 3) and the average regressions (columns 5 and 6) are very similar, implying that in the absence of high stakes, the

[37] We did not randomize incentives or stakes for questions in columns 4–6.

[38] In Table 1, we stacked the zero-sum and quintile ranks by question. So in column 1 of Table 1, the outcome variable is the household income and the regressors are the income quintile and zero-sum ranks, with a fixed effect for ranking type. Notice that the outcome variable is the same (household income) whether the regressor is a quintile or zero-sum ranking.

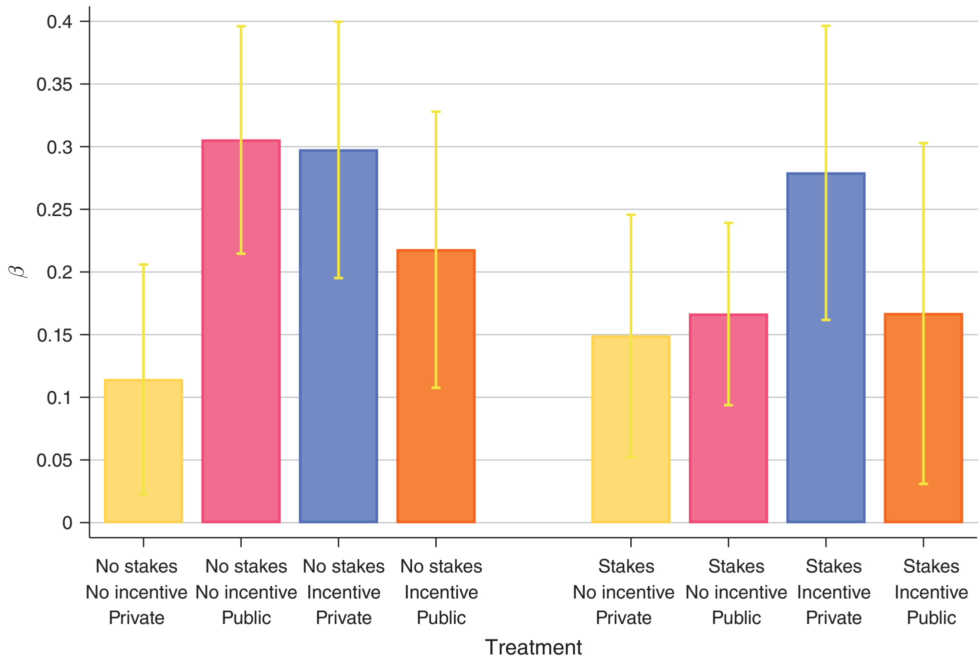

FIGURE 4. CORRELATION BETWEEN AVERAGE RANK AND OUTCOME

Notes: Each bar plots the mean and the standard error of the correlation between percentile of average rank (excluding the self-rank) and percentile of the outcome. Each bar corresponds to one of the eight treatment statuses, described below each bar. For a description of how the outcome and the average rank are constructed, see specification (10). In Table 6, we test the differences between these treatment conditions in a regression framework.

value of information from relative and quintile ranks is very similar. We also cannot reject that respondents misreport by the same amount in either type reporting method.

Overall, we find that in the presence of real stakes, misreporting is an important problem.

F. *Can Mechanism Design Tools Improve the Accuracy of Reports?*

Monetary Incentives and Public Reporting.—Can we use tools from mechanism design to generate incentives for truthful reporting? And, are these tools effective even in high stakes settings? We test the efficacy of two tools: payments for the accuracy of reports and reporting in public versus private.[39] Because we cross-randomized the *Public*, *Payments*, and *Stakes* treatments, there are eight treatment combinations in which we can evaluate the accuracy of reports. The average accuracy of reports in each treatment cell is depicted in Figure 4. In Table 6, we provide

[39] We discuss the mechanics of our payment rule in Sections IIB and online Appendix C.

TABLE 6—HOW DO INCENTIVES AND PUBLIC REPORTING AFFECT RESPONSES?

	Pooled questions (1)	Pooled questions (2)	Pooled questions (3)	Pooled questions (4)
Average rank	0.205	0.163	0.147	0.119
	(0.041)	(0.049)	(0.055)	(0.058)
Average rank × public	0.008	−0.031	0.196	0.019
	(0.060)	(0.067)	(0.076)	(0.070)
Average rank × incentives	−0.029	−0.078	0.129	0.143
	(0.063)	(0.070)	(0.074)	(0.081)
Average rank × incentives × public	−0.022	0.063	−0.262	−0.119
	(0.095)	(0.104)	(0.102)	(0.107)
Who is ranked?	Not self	Not self	Not self	Not self
Treatment	[No stakes]	[Stakes]	[No stakes]	[Stakes]
Observations	2,834	2,893	2,846	2,902
Number of households	1,339	1,339	1,345	1,345

Notes: Specification: This table estimates specification (10) in the paper. The regressions include incentives, public, and incentives × public, but the coefficients are not reported in the table. Average rank in columns 1 and 2 is the percentile of the rank that an entrepreneur assigns to herself on a particular question. In columns 3 and 4, average rank is the percentile of the average ranking the entrepreneur was given by her peers for a particular question (excluding the rank she assigned to herself). The unit of observation is the rankee by question. In columns 1 and 3, we limit the analysis to the *NoStakes* treatment group. In columns 2 and 4, we limit the analysis to the *Stakes* group. All regressions include ranking question, randomization strata, survey month, and surveyor fixed effects.
Outcome variables: We pool across questions 1–3 in panel A of Table 1 (in order to be comparable across questions, the outcome variable is percentilized) and that is the outcome across all columns of the table. For a description of the data that produced the outcome variables, see online Appendix D.

quantitative evidence of the *Public* and *Payments* treatments on the accuracy of reports. Again, following specification (2) we estimate,

$$(10)\quad Y_{ijq} = \eta_0 + \eta_1 \overline{Rank}_{ijq} + \eta_2 Public_j \times \overline{Rank}_{ijq} + \eta_3 Payments_j \times \overline{Rank}_{ijq} + \eta_4 Public_j \times Payments_j \times \overline{Rank}_{ijq} + \eta_5 Public_j + \eta_6 Payments_j + \eta_7 Public_j \times Payments_j + \gamma_r + \theta_m + \tau_s + \delta_q + \epsilon_{ijq}.$$

The coefficient η_1 identifies the accuracy of reports in groups in which respondents do not receive incentive payments and report in private. The coefficients on the first three interaction terms identify the additional accuracy due to reporting (i) in public without monetary payments (η_2), (ii) in private with monetary payments (η_3), and (iii) in public with monetary payments (η_4).[40]

To determine how these tools perform in a high stakes setting, we split results by *NoStakes* (odd columns) and *Stakes* (even columns). We also split the results by whether a respondent is reporting about herself (columns 1 and 2) or about her peers (columns 3 and 4). We find that community members are both more

[40]To reduce clutter in the regression tables, we have omitted the coefficients $Public_j$, $Payments_j$, $Public_j \times Payments_j$ from Table 6 as they do not contain information relevant for the interpretation of results, but rather simply adjust the intercept.

accurate and less responsive to incentives for truthfulness when reporting about themselves. Shifting from the no stakes (column 1) to the high stakes (column 2) setting decreases the accuracy of self-reports by 20.5 percent, although this difference is not statistically significant. Moreover, neither payments for truthfulness nor public reporting have any impact on the accuracy of self-reports. Note, though, that the accuracy of their self-reports (0.16 [standard error = 0.05] in column 2) in the high stakes setting is approximately the same as the accuracy of reports about others in the group in the private and no payments treatment (0.15 [standard error = 0.06] in column 3).

When reporting about others, incentives for truthfulness can have a large impact on respondents' accuracy. First, in the *NoStakes* setting, the *Payments* and *Public* treatments both double the accuracy of reports (they each lead to increase in accuracy between 0.13 [standard error = 0.07] and 0.20 [standard error = 0.08]). The coefficient on the treatment in which respondents receive monetary incentives and report in public is large and negative (*AverageRank* × *Payments* × *Public*). But, we can reject at the 10 percent level that the accuracy of information in this group is the same as in the private reporting and no monetary incentives group. We therefore interpret the negative coefficient as an indication that monetary payments and public reporting are substitutes.

In the *Stakes* setting, we find that the *Public* treatment no longer has a significant impact on accuracy. The impact of public reporting on accuracy is ambiguous ex ante. There may be pressure for respondents to up-rank their family members, but there may also be pressure from nonfamily members and other peers to be truthful. When we introduce stakes, both of these pressures are intensified. That we find different impacts of observability in the *Stakes* and *NoStakes* treatment might reflect the differing intensities of these two competing forces, or it might reflect a lack of precision in our estimates. That the *Public* treatment has no impact in the *Stakes* setting may cast doubt on its usefulness as a tool to induce truthfulness in practice.

The monetary payments treatment is still effective when allocation of resources is at stake: when reports are made in private, the *Payments* treatment improves accuracy by 0.14 (standard error = 0.08), which is an increase in accuracy of over 100 percent. Once again, the interaction between *Payments* and *Public* is large and negative, which may imply that these are substitutes or may indicate that payments are less effective in public.

Finally, in online Appendix Table A20 we examine the noninteracted impact of the *Public* treatment and the *Payments* treatment. While on average the *Public* treatment has no effect on accuracy, the *Payments* treatment improves accuracy in the *Stakes* treatment by 0.09 (standard error = 0.05), though it narrowly misses statistical significance at traditional levels. On average, the *Payments* treatment almost entirely corrects the misreporting attributable to the *Stakes* treatment.

Our results thus far present a trade-off. Community members are most informed about themselves, but their reports about themselves are also unresponsive to incentives for accuracy. In contrast, when reporting about others, while community members may have lower baseline information, they are also more responsive to monetary incentives. Therefore, which of these sources of information is more valuable may depend on the incentives community members face to distort reports.

How Do Respondents Distort Their Reports?—So far we have established that respondents distort their reports when the distribution of resources is at stake, and that simple mechanisms can realign incentives for accuracy. Lastly we ask, in whose favor do respondents distort their reports? At the start of the ranking exercise, we asked respondents to report their relationship with each peer in the group. We also asked each respondent to identify each other person's closest peer in the group. An entrepreneur's *cross-reported peer* is the peer that is most frequently reported as their closest friend in the group.

To assess which people are favored by respondents' lies, we analyze how the rankings themselves (not just accuracy) are affected by proximity between peers. Specifically we estimate

$$(11) \quad Rank_{hijq} = \eta_0 + \eta_1 Characteristic_{hij} + \eta_2 Public_j \times Characteristic_{hij} + \eta_3 Payments_j \times Characteristic_{hij} + \eta_4 Public_j \times Payments_j \times Characteristic_{hij} + \eta_5 Public_j + \eta_6 Payments_j + \eta_7 Public_j \times Payments_j + \gamma_n + \theta_m + \tau_s + \delta_q + \epsilon_{hijq},$$

where $Rank_{hijq}$ is the rank that entrepreneur h assigns to entrepreneur i in group j for question q, $Characteristic_{hij}$ is a dummy for the relationship between entrepreneurs h and i in group j, and the rest of the variables are defined as above. Results are presented in online Appendix Table A21. We see that respondents up-rank family members and cross-reported peers relative to other peers in the group in the absence of incentives and in private. But incentives and publicity reduce the average rank assigned to either of these groups.

To some extent, respondents may also have been distorting their reports to favor poorer members of their groups. We implemented a regressive policy that targets grants to members of the community that have the highest income, assets, or profits, and respondents may have attempted to instead target the grants to those most in need.[41] It is difficult to test for this explicitly, as there is a mechanical relationship between the fact that reports are less accurate in the high stakes condition and the fact that poorer respondents were ranked more highly. In any event, our results that respondents are more likely to up-rank their family, friends, and cross-reported peers indicates that a desire for redistribution to the needy cannot account for the full extent of misreporting.

Identifying the Most Informed Community Members.—Next, we provide investigate whether the most informed members of the community can be identified ex ante. Recall, we asked respondents to name the person who would provide the

[41] We note that in settings where lenders or governments want to target the most able entrepreneurs, a regressive transfer may be justified.

most accurate reports on average. To assess whether this exercise was a success, we estimate

$$(12)\qquad Y_{ijkq} = \beta_0 + \beta_1 Rank_{ijkq} + \beta_2 MostInformed_{ijk} + \beta_3 Rank_{ijkq} \times MostInformed_{ijk} + \delta_q + \gamma_r + \theta_m + \tau_s + \epsilon_{ijkq}.$$

Observations are at the ranker-rankee-question level and $MostInformed_{ijk}$ is a dummy for whether a majority of the group selected ranker k as the most informed reporter. To increase power, we stack the percentilized outcomes and ranks across all of the columns presented in panel A of Table 1, and add a question fixed effect (δ_q) to the regression model. Results are presented in online Appendix Table A22. We do not find evidence that people who are selected as the most informed provide more accurate information than other group members.

Finally, we investigate whether entrepreneurs are especially informed about those who are like themselves. Specifically we estimate

$$(13)\qquad Y_{ijkq} = \beta_0 + \beta_1 Rank_{ijkq} + \beta_2 Characteristic_{ijk} + \beta_3 Rank_{ijkq} \times Characteristic_{ijk} + \delta_q + \gamma_r + \theta_m + \tau_s + \epsilon_{ijkq}.$$

$Characteristic_{ijk}$ is a dummy for whether ranker k and rankee i share a characteristic in common, and all other features are as in specification (12). We restrict attention to outcomes that correspond to the individual entrepreneur (profits, digit span, and hours worked). Results are presented in online Appendix Table A23. Column 1 restricts the sample to male entrepreneurs, and column 2 restricts attention to female entrepreneurs. In both columns the characteristic of interest is the ranker's gender. In the next three columns we restrict the sample to entrepreneurs from each of the top three industries in our study (tailors, vegetable vendors, and *kirana* shops) and the characteristic of interest is whether the entrepreneurs are in the same industry. While β_3 is not statistically significant across any of the columns, the results suggest that women may have an advantage in ranking women (column 2), while there is no evidence that men or women have an advantage in ranking men. Further, kirana shop owners may have an advantage in ranking other kirana shop owners. In other industries we find no evidence of a within-industry advantage in ranking entrepreneurs.

V. Discussion

We find that community members have information about their peers that is valuable for targeting even after controlling for a wide range of observable characteristics. Not only can community members identify characteristics of their peers' enterprises, they can also predict which of their peers have high returns to capital. But community information is also susceptible to strategic misreporting. In particular, we identify a tendency for respondents to favor their friends and family members in their reports. Moreover, misreporting is exacerbated when respondents are told that their reports will influence the distribution of grants. However we also find that a variety of techniques motivated by mechanism design theory are effective in realigning incentives for truthfulness. Relatively small monetary payments for accuracy and cross reporting techniques both substantially improve the accuracy of reports.

Is it worth it to invest in collecting community information and providing incentives to respondents? We calibrated the payment rule to pay, on average, 25 rupees per question per respondent. In total, we paid 17,000 rupees in incentives for the marginal returns question. If a lender were distributing 450 loans (as we did with grants), this would increase the cost on each loan by approximately 40 rupees per month—far less than our estimates of the marginal returns to capital for entrepreneurs who fall in the top third of community ranks.[42]

Our hope is that the peer elicitation method identified in this paper can be useful for targeting in poorly developed financial markets in low-income countries, where information asymmetries are prevalent. Of course there are limits to the extent to which these tools can be immediately adopted by lenders or governments in such settings. For instance, respondents in our study were not given time to learn about and adapt to the payment rule. We cannot be certain whether respondents' predictions about one another's productivity reflect fixed or time-varying enterprise characteristics. And even if lenders can identify high-productivity entrepreneurs, lenders may not be able to extract borrowers' marginal profits attributable to a loan. These limitations represent fertile areas for future research. Even so, we expect that the broad conclusions of this study remain useful to interested parties. Community members are well informed about one another's marginal return to capital; incentives to misreport information in favor of friends and family loom large in high stakes situations; simple techniques like monetary payments for accuracy are effective at realigning incentives.

The tools developed in this paper may also prove useful in other contexts in which researchers and policymakers aim to target resources using community information. This may be especially true when targeting is to be done based on *treatment effects* rather than observable characteristics, and in settings where the incentives of community members and policymakers may not be fully aligned.

REFERENCES

Alatas, Vivi, Abhijit Banerjee, Rema Hanna, Benjamin A. Olken, and Julia Tobias. 2012. "Targeting the Poor: Evidence from a Field Experiment in Indonesia." *American Economic Review* 102 (4): 1206–40.

Alatas, Vivi, Abhijit Banerjee, Rema Hanna, Benjamin A. Olken, Ririn Purnamasari, and Matthew Wai-Poi. 2019. "Does Elite Capture Matter? Local Elites and Targeted Welfare Programs in Indonesia." *AEA Papers and Proceedings* 109: 334–39.

Austin, Peter. 2011. "An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies." *Multivariate Behavioral Research* 46 (3): 399–424.

Banerjee, Abhijit, Dean Karlan, and Jonathan Zinman. 2015. "Six Randomized Evaluations of Microcredit: Introduction and Further Steps." *American Economic Journal: Applied Economics* 7 (1): 1–21.

Barboni, Giorgia, and Parul Agarwal. 2020. "Knowing What's Good for You: Can a Repayment Flexibility Option in Microfinance Contracts Improve Repayment Rates and Business Outcomes?" Unpublished.

[42] As it took about 20 minutes to elicit the marginal returns rankings, the associated cost of labor is negligible.

Basurto, Maria Pia, Pascaline Dupas, and Jonathan Robinson. 2019. "Decentralization and Efficiency of Subsidy Targeting: Evidence from Chiefs in Rural Malawi." *Journal of Public Economics*. https://doi.org/10.1016/j.jpubeco.2019.07.006.

Baum, J. Robert, and Edwin A. Locke. 2004. "The Relationship of Entrepreneurial Traits, Skill, and Motivation to Subsequent Venture Growth." *Journal of Applied Psychology* 89 (4): 587–98.

Beaman, Lori, and Jeremy Magruder. 2012. "Who Gets the Job Referral? Evidence from a Social Networks Experiment." *American Economic Review* 102 (7): 3574–93.

Beaman, Lori, Dean Karlan, Bram Thuysbaert, and Christopher Udry. 2020. "Selection into Credit Markets: Evidence from Agriculture in Mali." Unpublished.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen. 2014. "Inference on Treatment Effects after Selection among High-Dimensional Controls." *Review of Economic Studies* 81 (2): 608–50.

Bernhardt, Arielle, Erica Field, Rohini Pande, and Natalia Rigol. 2019. "Household Matters: Revisiting the Returns to Capital among Female Microentrepreneurs." *American Economic Review: Insights* 1 (2): 141–60.

Bloch, Francis, and Matthew Olckers. 2019. "Friend Based Ranking." Unpublished.

Bluedorn, Allen C., Thomas J. Kalliath, Michael J. Strube, and Gregg D. Martin. 1999. "Polychronicity and the Inventory of Polychronic Values (IPV): The Development of an Instrument to Measure a Fundamental Dimension of Organizational Culture." *Journal of Managerial Psychology* 14 (3–4): 205–31.

Bryan, Gharad, Dean Karlan, and Jonathan Zinman. 2015. "Referrals: Peer Screening and Enforcement in a Consumer Credit Field Experiment." *American Economic Journal: Microeconomics* 7 (3): 174–204.

Chodorow-Reich, Gabriel, Gita Gopinath, Prachi Mishra, and Abhinav Narayanan. 2020. "Cash and the Economy: Evidence from India's Demonetization." *Quarterly Journal of Economics* 135 (1): 57–103.

Convergences. 2019. *Microfinance Barometer 2019*. Paris: Convergences.

Dal Bó, Ernesto, Frederico Finan, Nicholas Y. Li, and Laura Schechter. 2018. "Government Decentralization under Changing State Capacity: Experimental Evidence from Paraguay." Unpublished.

Danz, David, Lise Vesterlund, and Alistair J. Wilson. 2020. "Belief Elicitation: Limiting Truth Telling with Information on Incentives." Unpublished.

DellaVigna, Stefano, and Devin Pope. 2018. "Predicting Experimental Results: Who Knows What?" *Journal of Political Economy* 126 (6): 2410–56.

de Mel, Suresh, David J. McKenzie, and Christopher Woodruff. 2009. "Measuring Microenterprise Profits: Must We Ask How the Sausage is Made?" *Journal of Development Economics* 88 (1): 19–31.

de Mel, Suresh, David McKenzie, and Christopher Woodruff. 2008. "Returns to Capital in Microenterprises: Evidence from a Field Experiment." *Quarterly Journal of Economics* 123 (4): 1329–72.

Fafchamps, Marcel, and Christopher Woodruff. 2017. "Identifying Gazelles: Expert Panels vs. Surveys as a Means to Identify Firms with Rapid Growth Potential." *World Bank Economic Review* 31 (3): 670–86.

Fafchamps, Marcel, David McKenzie, Simon Quinn, and Christopher Woodruff. 2014. "Microenterprise Growth and the Flypaper Effect: Evidence from a Randomized Experiment in Ghana." *Journal of Development Economics* 106: 211–26.

Giné, Xavier, and Dean S. Karlan. 2014. "Group versus Individual Liability: Short and Long Term Evidence from Philippine Microcredit Lending Groups." *Journal Of Development Economics* 107: 65–83.

Honorati, Maddalena, Ugo Gentilini, and Ruslan G. Yemtsov. 2015. *The State of Social Safety Nets 2015*. Washington, DC: World Bank Group.

Hussam, Reshmaan, Natalia Rigol, and Benjamin N. Roth. 2016. "Knowing Thy Neighbor: What Information Neighbors Have and How Best to Elicit It." AEA RCT Registry, March 7. https://doi.org/10.1257/rct.1109-2.0.

Hussam, Reshmaan, Natalia Rigol, and Benjamin N. Roth. 2022. "Replication Data for: Targeting High Ability Entrepreneurs Using Community Information: Mechanism Design in the Field." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/E151841V1.

Jayachandran, Seema. 2020. "Microentrepreneurship in Developing Countries." Unpublished.

Karlan, Dean, Adam Osman, and Jonathan Zinman. 2016. "Follow the Money Not the Cash: Comparing Methods for Identifying Consumption and Investment Responses to a Liquidity Shock." *Journal of Development Economics* 121: 11–23.

Klinger, Bailey, Asim Ijaz Khwaja, and Carlos del Carpio. 2013. *Enterprising Psychometrics and Poverty Reduction*. New York: Springer.

Maitra, Pushkar, Sandip Mitra, Dilip Mookherjee, Alberto Motta, and Sujata Visaria. 2017. "Financing Smallholder Agriculture: An Experiment with Agent-Intermediated Microloans in India." *Journal of Development Economics* 127: 306–37.

Maskin, Eric. 1999. "Nash Equilibrium and Welfare Optimality." *Review of Economic Studies* 66 (1): 23–38.

McClelland, David C. 1985. "How Motives, Skills, and Values Determine What People Do." *American Psychologist* 40 (7): 812–25.

McKenzie, David, and Christopher Woodruff. 2008. "Experimental Evidence on Returns to Capital and Access to Finance in Mexico." *World Bank Economic Review* 22 (3): 457–82.

McKenzie, David, and Dario Sansone. 2019. "Predicting Entrepreneurial Success Is Hard: Evidence from a Business Plan Competition in Nigeria." *Journal of Development Economics* 141: 102369.

Rigol, Natalia, and Benjamin N. Roth. 2020. "Loan Officer Incentives and Graduation from Microfinance: Evidence from Chile." Unpublished.

Rigol, Natalia, and Benjamin N. Roth. 2017. "Paying for the Truth: The Efficacy of Peer Prediction in the Field." Unpublished.

Rosenbaum, Paul R. 1987. "Model-Based Direct Adjustment." *Journal of the American Statistical Association* 82 (398): 387–94.

Rotter, Julian B. 1966. "Generalized Expectancies for Internal versus External Control of Reinforcement." *Psychological Monographs: General and Applied* 80 (1): 1–28.

Wager, Stefan, and Susan Athey. 2018. "Estimation and Inference of Heterogeneous Treatment Effects Using Random Forests." *Journal of the American Statistical Association* 113 (523): 1228–42.

Witkowski, Jens, and David Parkes. 2012. "A Robust Bayesian Truth Serum for Small Populations." Paper presented at the Proceedings of the 26th AAAI Conference on Artificial Intelligence, Toronto, Ontario, Canada, July 22.

This article has been cited by:

1. Mark York, Munther Dahleh, David C. Parkes. Eliciting Social Knowledge for Creditworthiness Assessment 428-445. [Crossref]