Identifying and Spurring High-Growth Entrepreneurship:

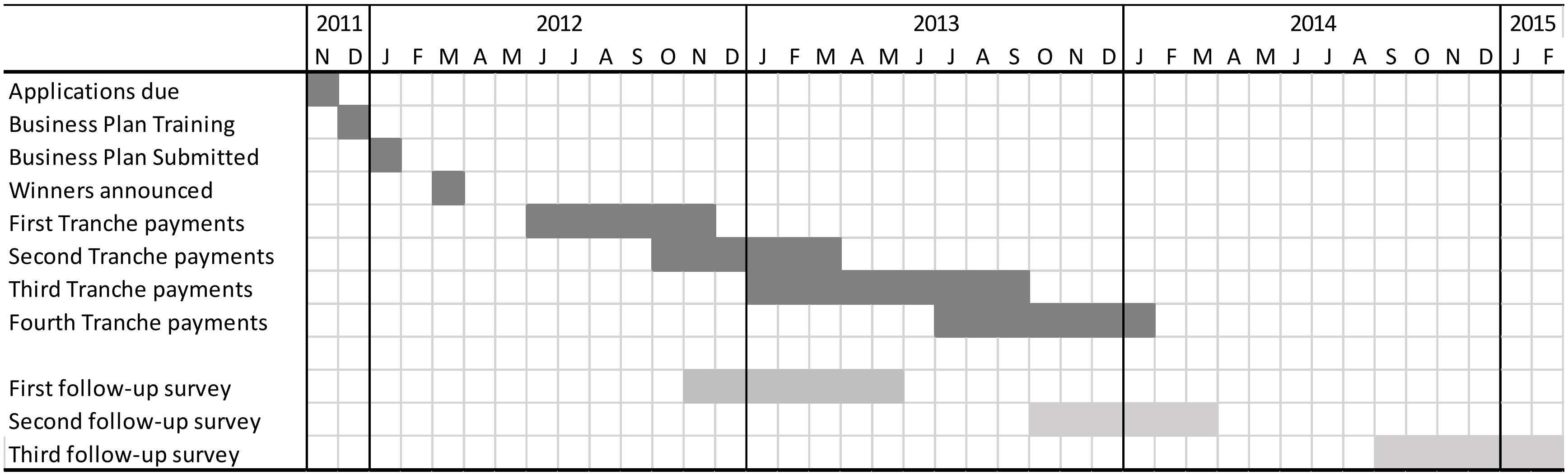
Experimental Evidence from a Business Plan Competition

David McKenzie, *World Bank*

APPENDICES (ONLINE ONLY)

List of Appendices

Appendix 1: Timeline



Note: long-term follow-up survey took place July-November 2016.

Appendix 2: Additional Details on Business Plan Competition

A2.1 Launch and Outreach

The program was advertised throughout the country over different television and radio stations. Adverts were also published in the newspapers with the widest coverage. Road shows were organized by the Ministry of Youth Development and private vendors in major cities of each geo-political area of the country targeting areas with large numbers of youth eligible for the competition. The six geo-political zones are: North-Central (includes Abuja), North-Eastern (includes Borno), North-Western (includes Kaduna), South-Eastern (includes Abia), South-South (includes Rivers and Delta), and South-Western (includes Lagos). Small and Medium Enterprise outreach events were also held in Lagos and Abuja.

A2.2 Information required in initial application

All applicants had to provide the following information:

- A statement as to why they want to be an entrepreneur, how they got their business idea, and why they will succeed.
- A description of their business idea, why it is innovative, what their market will be, why people will buy their products, who their competition is, how the business will make money, and the risks foreseen and how they will overcome them.

New applicants also had to provide:

- What the key steps needed to start the business are
- Description of their qualifications and experience
- How much money they need to start the business.

Existing business owners needed to provide information on:

- Years of operation, turnover, employment levels, and registration certificate number (firms did not need to be registered to apply, but if they won would need to register in order to be eligible to receive a grant).
- How much money they need to expand the business, how many more people they will employ if they do so, and what the projected annual turnover will be.

A2.3 Who Applied?

Nigeria has approximately 50 million youth aged 18 to 40. The almost 24,000 applications therefore represent only 0.05% of the overall youth population. Applicants are older on average, and more educated than the average Nigerian youth. Among the overall youth population, 5.5% have university education, compared to 52% of new business applicants and 54% of existing business applicants. Geographically we see that a higher proportion of youth applied from the

North-Central region (where Abuja is located) and the South-Western region (where Lagos is located), while the North-Western region had the lowest proportion of youth.

A2.4 Content of the Four-Day Business Plan Training

Training was run by EDC with support from Plymouth Business School, and was a four day course. The goal was to provide tools and techniques that would help both in writing a business plan and in running the business. The course covered:

- The different sections of what should go into a business plan – and what sort of things funders would look for in each section
- How to find out more about the competition and competitive environment; understanding your competitors and how you can differentiate yourself
- Business plan financials – putting together a balance sheet, cash flow forecast, and profit and loss forecast; financial planning and breakeven analysis.
- Legal and regulatory matters: different forms of legal registration and how to register, different forms of business (e.g. sole proprietorship, partnership, different company types), taxation responsibilities.
- Introduction to marketing strategy – creating a marketing plan, different strategies for selling, marketing research, market segmentation.
- Establishing an online presence and engaging with customers through social media
- Presentation skills and developing a funding pitch and sales pitch
- Strategies for growth – the role of horizontal and vertical integration, of product diversification, and of Strategic Alliances.
- A quick introduction to the IFC-developed SME Toolkit available online, and all participants were also given a CD copy of this.

A2.5 Information Collected at Time of Submission of Business Plan

The business plan collected information on the business profile including the product, its customer base, pricing, the experience and qualifications of the owner, a detailed description of physical capital, premises, form of business organization, cash flow and projected income statements for the first year, financing strategy including other sources of capital, use of e-commerce, marketing plans, and the perceived increase in employment from getting a grant from the YouWiN!! Project. In addition to this business plan, a baseline data sheet was collected which asked about previous business courses taken, current financing, demographic characteristics of the owner, time spent abroad, risk attitudes, reasons for wanting to own a business rather than work in salary work, self-assessed entrepreneurial efficacy, household asset ownership, and follow-up information to enable re-surveying in the future.

A2.6 Scoring of Business Plans

The business plan was scored out of 100, using the following scoring scheme:

- Articulation of market potential (10 points) – this had subcategory points for describing the existing business environment, what the gap in the market was, the existence of substitutes and competitors, and what the potential customer demand would be.
- Time to market (5 points) – this had subcategory points for the product channels, and for the product delivery time.
- Understanding of the industry (10 points) – this had subcategory points for describing the key stakeholders in the industry, the industry value chain, and for SWOT (strengths, weaknesses, opportunities, and threats) analysis of the industry.
- Job creation (25 points) – this awarded points for the potential of the firm to create employment.
- Financial viability (5 points) – this awarded points for the existence of a profit and loss statement and balance sheet, for having a cash flow statement, and for the accuracy and reasonableness of Figures.
- Financing sources (5 points) – this awarded points for personal funding sources, for working capital, and for how much they would rely on YouWin.
- Financial sustainability (10 points) – this awarded points for profitability, liquidity, the growth trend, and for being environmentally friendly.
- Ability to manage the business (10 points) – this awarded points for being an existing business, for the qualifications of the owner, for the managerial and technical expertise of the owner, for the business organization, and for the types of controls in place.
- Risk assessment/Mitigation (10 points) – this awarded points for assessing risks facing the business growth and for the planned mitigation strategies to address them.
- Capstone score (10 points) – this was a final category where the scorer could assign points based on their overall assessment and comfort level in the business after reading the business plan.

A2.7 Quality Checking and Finalization of Winners

After the 1,200 provisional winners were selected, a DFID-procured firm (Growbridge Advisors, supported by Nigerian consultants) reviewed all winning business plans to validate whether the award amount asked for was reasonable given the proposal, and to propose business milestones and targets, along with a disbursement schedule. As a result of this process, 18 of the original 1,200 winners (3 national, 2 zonal and 13 ordinary merit winners) were disqualified based on an assessment that they required significantly more than 10 million Naira for their business, or that their financial projections were unrealistic. These 18 disqualified proposals were replaced with 18 businesses from the ordinary winner control group. 9 of these replacements were randomly chosen from the same regions and new/existing business status as the firms they were replacing. However, given the rapid finalization of the winners in time for an official announcement and the short time frame for assessing disqualifications, there was a need for 9 further replacements during a day in which the author was on an airplane. These other 9 replacements were chosen as the highest scoring ordinary winner control group in the zones that they were replacing.

Appendix 3: Amounts Received by Winners and Timing Relative to Surveys

Table A3: Summary of Amounts Received by Winners

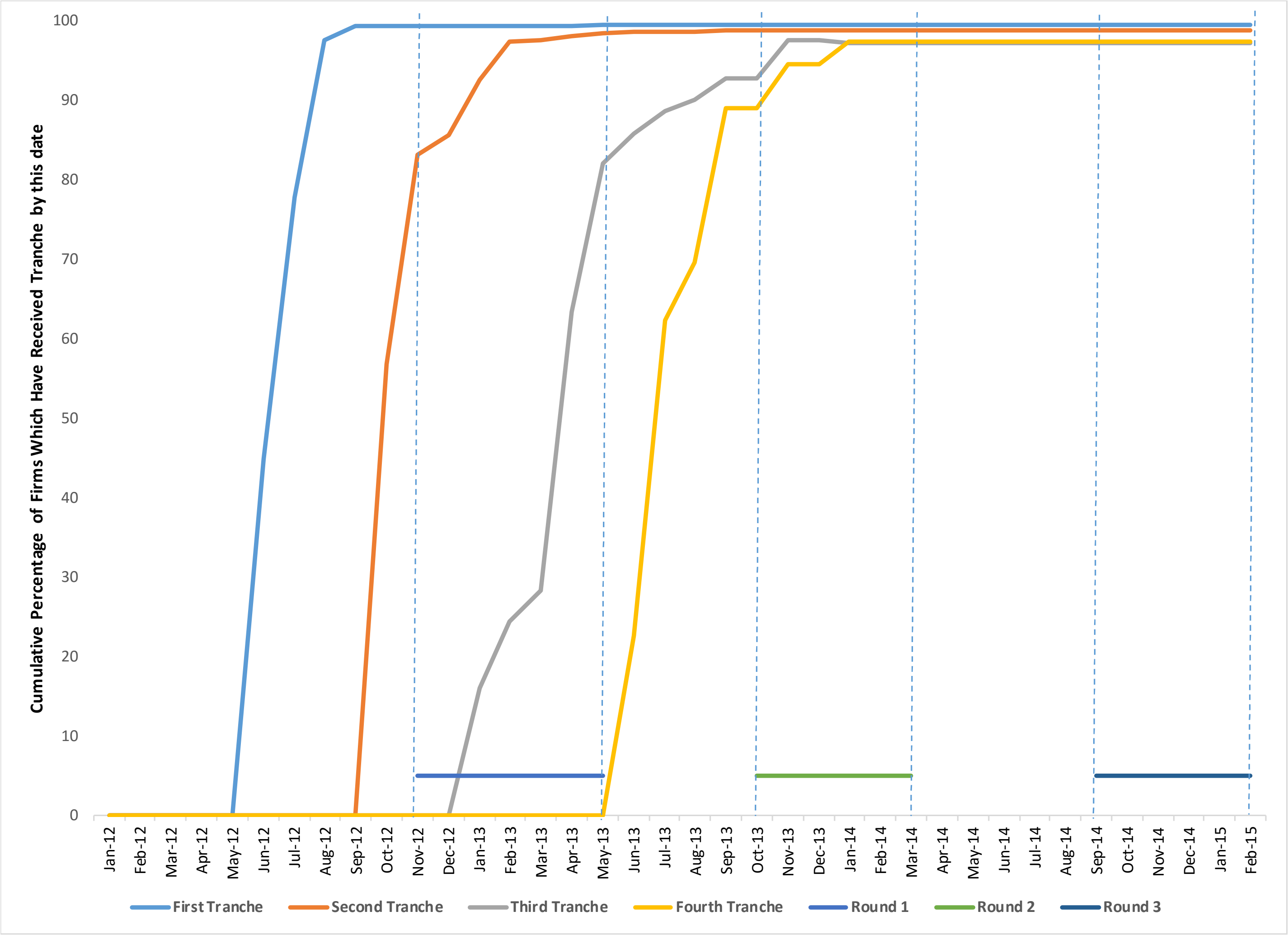
	Mean	S.D.	10th	Median	90th	Max
Amount Received as First Tranche						
Naira	1079023	768276	400000	1000000	2268452	5000000
USD	6873	4893	2548	6369	14449	31847
Amount Received by First Follow-up Survey						
Naira	3591152	1820305	750000	4000000	5000000	8000000
USD	22874	11594	4777	25478	31847	50955
Total Amount Received over all Tranches						
Naira	7691604	2754758	3400000	9000000	10000000	10000000
USD	48991	17546	21656	57325	63694	63694
Months since First Tranche (Follow-up 1)	4.8	1.9	3	5	7	10
Months since First Tranche (Follow-up 2)	15.4	1.6	13	15	17	20
Months since First Tranche (Follow-up 3)	27.1	1.4	25	27	29	31
Months since Last Tranche (Follow-up 3)	14.4	1.5	12	14	16	18

Note: Exchange rate of 157 Naira = 1 USD used. Data for the 1200 winners.

Table A3 shows the amounts awarded for the pooled group of 1200 winners. The experimental winners had a mean grant of $46,627 and the non-experimental winners a mean grant of $52,952. Among the experimental winners, the average grant amount was similar for the new firms and existing firms ($46,632 vs $46,619).

Figure A3 shows the share of firms which have received each of the four tranches over time, and related this to survey timing. We see the first survey occurs when most firms have two tranches, the second when most firms have all four tranches, and the third survey round over a year after most firms have received all tranches. The long-term follow-up round is not shown, but corresponds to more than three years after all funding has been received.

Figure A3: Cumulative Percentage of Firms Which Have Received Each Tranche and Survey Timing



Appendix 4: Business Sectors Proposed by Winners

Table A4: Business Sectors Proposed by Winners

	Existing Firms				New Firms			
	National Winners	Zonal Winners	Ordinary Winners	All Winners	National Winners	Zonal Winners	Ordinary Winners	All Winners
Retail trade	4.0	4.5	4.2	4.3	4.1	0	4.9	4.4
Food preparation or restaurant	4.0	3.0	2.5	3.2	2.7	6.8	7.2	6.5
Personal services	1.4	4.5	5.1	3.6	1.4	0	5.3	4.4
Tailoring/dressmaking/shows	4.5	7.5	5.4	5.5	5.4	2.3	3.1	3.4
Furniture manufacturing	0.5	0.8	1.8	1.1	0	0	0.7	0.5
Crafts (masks, jewellery, etc.)	0.5	0.8	2.5	1.4	0	2.3	1.1	1.1
Other manufacturing	17.5	11.2	13.4	14.4	14.9	15.9	13	13.4
Repair services	1.0	0.8	2.1	1.4	1.4	0	0.9	0.9
IT and Computer services	18.8	14.9	17.3	17.4	5.4	0	7.8	6.9
Accounting, legal, and medical services	3.6	1.5	1.4	2.2	5.4	0	2.5	2.7
Other professional services	13.5	7.5	7.6	9.6	10.8	13.6	6.5	7.6
Transportation	0.9	0.8	1.8	1.3	6.8	0	1.4	1.9
Construction work	2.3	6.7	4.8	4.3	1.4	2.3	2.7	2.5
Agricultural production	17.0	25.4	20.2	20.2	27	40.9	32.4	32.3
Other industries	10.8	9.7	9.4	9.9	13.5	15.9	10.5	11.3

Appendix 5: Propensity Score Matching Estimates of the Treatment Effects for Non-Experimental Winners

The experimental estimation gives the impact of winning the program (and being assigned a large grant) for the semi-finalists in the experimental pool. To measure the impact of winning on the national and zonal winners (hereafter non-experimental winners), I use matching, comparing these winners to the experimental control group.

The follow-up surveys attempted to reach the 475 national and zonal winners (which excludes five winners who were disqualified). The response rates were very similar to those of the experimental winners, as shown in Appendix Table A5.1:

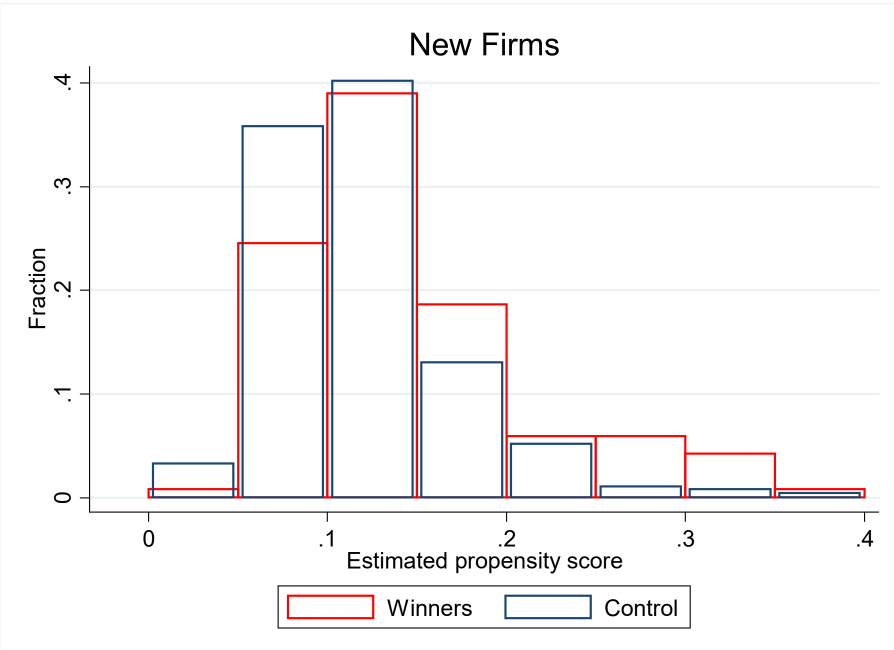
Appendix Table A5.1: Survey Response Rates for Non-Experimental Winners Compared to Response Rates for Winners

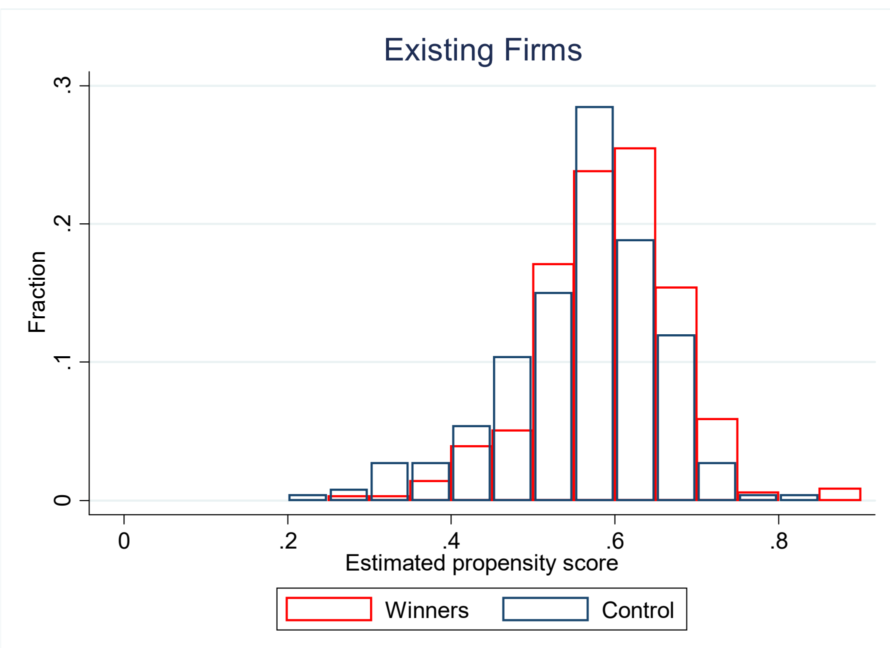
	New Firms			Existing Firms		
	First Survey	Second Survey	Third Survey	First Survey	Second Survey	Third Survey
Panel A: Information available on whether or not they operate a firm						
Experimental Winners	0.840	0.922	0.869	0.856	0.942	0.910
National and Zonal Winners	0.864	0.873	0.831	0.824	0.941	0.894
Panel B: Responded to the Survey						
Experimental Winners	0.796	0.900	0.847	0.838	0.939	0.906
National and Zonal Winners	0.831	0.856	0.805	0.798	0.938	0.891

The pre-analysis plan pre-specified and pre-coded propensity score matching on the basis of the variables used for the balance check in Table 1, with the exception of the first round application mark and business plan score.[1] It therefore matches on gender, age, marital status, education, international migration experience, risk attitude, household wealth proxies, and the type of sector the applicant proposes having a business in. In addition the existing businesses are also matched on the number of workers they had at the time of application, and whether they had ever had a formal loan at the time of application. Figure A5 shows the propensity score distributions overlap well for the non-experimental winners and control groups. The balancing property is satisfied for both new and existing firms.

[1] One of the referees suggested that I also match on the application and business plan score. The problem with this is that these scores predict very well who becomes a zonal and national winner, so that there is very little overlap in the propensity score distributions, and the distributions are sparse within the common support. As a result the propensity score balancing property is not satisfied, and estimates using these scores are unreliable.

Figure A5: Propensity Score Distributions for Non-Experimental Winners and Controls Using Pre-specified Propensity Scores





The propensity score estimates are then obtained using kernel-based matching with a Gaussian kernel within the common support, with bootstrapped standard errors.

For robustness I compare these propensity score estimates to two alternatives, inspired by referee comments. The first is the bias-adjusted nearest-neighbor matching approach of Abadie et al. (2004) using the same set of covariates as for the propensity-score matching.[2] This approach matches on the basis of Mahalnobis distance between a set of covariates, rather than reducing them first to the propensity score and then matching on this single variable. The advantage of this approach is that it then more easily allows for exact matching on particular variables. In particular, since zonal winners were chosen separately within each of the six regions of Nigeria, I consider an estimate which requires exact matching within region (separately for new and existing firms), and then nearest-neighbor matching within these regions.

The conditions are reasonably promising for matching to be reliable, since both the non-experimental winners and control group self-selected into the program at the same point in time, had both already survived screening on the initial application, and to get to the semi-finals stage, and have similar observable backgrounds. However, I do not have multiple periods of pre-application data to match on growth trajectories for the existing firms, and since the winners were judged to have better growth prospects than the control group, matching may deliver an upper bound on the effectiveness of the program if the two groups differ in unobservable determinants of success.

The following two tables provide the matching estimates for new and existing firms respectively, focusing on the key outcomes of opening a business, total employment and exceeding a 10 worker threshold, the sales and profits index, and the innovation index. In general we see that the matching estimates are similar in magnitude and statistical significance across the three specifications.

The estimated treatment effects on operating a firm, reflecting start-up and survival, are similar in magnitude to the experimental estimates. This reflects the fact that almost all winners started firms and kept them going over all three years: 91.3 percent of the new firm experimental winners and 91.8 percent of the new firm non-experimental winners were operating at the time of the third follow-up; as were 95.6 percent of the existing firm experimental winners and 95.0 percent of the existing firm non-experimental winners.

In contrast, the matching estimates show larger impacts on employment creation and on sales and profits for the non-experimental winners than the experimental winners. However, we cannot say whether this larger impact reflects selection on unobservables (so that the very top scoring proposals are from applicants who would have grown larger firms even without the intervention) or whether it reflects the intervention being more successful for firms with the highest scores. Since analysis of heterogeneity in treatment effects does not find stronger effects for those with higher business plan scores amongst the semi-finals (section 6.5), if this result extrapolates outside of the experimental score-range, then this points to selection being the explanation. In contrast,

[2] Abadie, A., D. Drukker, J. L. Herr, and G. W. Imbens. 2004. Implementing matching estimators for average treatment effects in Stata. Stata Journal 4(3): 290-311.

the estimated impacts on the innovation index are very similar for the non-experimental winners as for the experimental winners.

Appendix Table A5.2: Matching Estimates of Impact for National and Zonal Winners for New Firms

	Sample Size	Pre-specified Propensity Score Matching	Nearest Neighbor	Nearest Neighbor within region	Experimental Estimate for Comparison
Operates a firm at time of survey					
Round 1	717	0.250***	0.255***	0.221***	0.213***
		(0.044)	(0.055)	(0.067)	(0.029)
Round 2	838	0.414***	0.394***	0.413***	0.358***
		(0.021)	(0.046)	(0.050)	(0.023)
Round 3	785	0.382***	0.393***	0.339***	0.373***
		(0.031)	(0.052)	(0.060)	(0.024)
Total Employment					
Round 1	719	2.707***	2.828***	2.715**	1.426*
		(0.785)	(1.074)	(1.286)	(0.732)
Round 2	843	11.162***	12.863***	13.060***	6.012***
		(2.280)	(2.279)	(2.862)	(0.412)
Round 3	748	7.007***	7.625***	7.037***	5.227***
		(1.059)	(1.141)	(1.173)	(0.469)
Firm of 10+ Workers					
Round 1	719	0.142***	0.159***	0.158***	0.024
		(0.044)	(0.046)	(0.055)	(0.020)
Round 2	843	0.344***	0.404***	0.374***	0.288***
		(0.047)	(0.050)	(0.061)	(0.026)
Round 3	748	0.261***	0.308***	0.287***	0.229***
		(0.055)	(0.054)	(0.061)	(0.028)
Aggregate Index of Sales and Profits					
Round 1	723	-0.023	-0.048	-0.149*	0.016
		(0.064)	(0.078)	(0.086)	(0.047)
Round 2	837	0.442***	0.501***	0.495***	0.298***
		(0.078)	(0.083)	(0.100)	(0.036)
Round 3	770	0.313***	0.350***	0.309***	0.167***
		(0.063)	(0.093)	(0.111)	(0.042)
Innovation Index					
Round 1	723	0.145***	0.167***	0.156***	0.099***
		(0.034)	(0.038)	(0.044)	(0.019)
Round 2	784	0.262***	0.293***	0.277***	0.270***
		(0.029)	(0.039)	(0.046)	(0.018)
Round 3	681	0.293***	0.278***	0.249***	0.219***
		(0.037)	(0.039)	(0.045)	(0.019)

Notes: Standard errors in parentheses, *, **, *** indicate significance at the 10, 5, and 1 percent levels respectively. Pre-specified propensity score matching uses kernel matching within the common support and bootstraps standard errors. Nearest neighbor matching uses the Abadie et al. (2004) matching approach.

Appendix Table A5.3: Matching Estimates of Impact for National and Zonal Winners for Existing Firms

	Sample Size	Pre-specified Propensity Score Matching	Nearest Neighbor	Nearest Neighbor within region	Experimental Estimate for Comparison
Operates a firm at time of survey					
Round 1	485	0.097***	0.128***	0.107***	0.082***
		(0.027)	(0.029)	(0.029)	(0.027)
Round 2	573	0.134***	0.125***	0.132***	0.130***
		(0.021)	(0.026)	(0.027)	(0.025)
Round 3	533	0.202***	0.195***	0.222***	0.196***
		(0.033)	(0.036)	(0.036)	(0.031)
Total Employment					
Round 1	471	3.149***	3.453***	3.957***	1.461*
		(0.777)	(0.977)	(1.066)	(0.808)
Round 2	571	5.739***	5.845***	7.094***	2.521*
		(1.569)	(1.731)	(1.310)	(1.366)
Round 3	528	7.338***	7.417***	7.413***	4.391***
		(0.808)	(0.991)	(0.982)	(0.674)
Firm of 10+ Workers					
Round 1	471	0.102**	0.099**	0.131***	0.055
		(0.043)	(0.050)	(0.048)	(0.041)
Round 2	571	0.347***	0.363***	0.389***	0.211***
		(0.041)	(0.047)	(0.044)	(0.041)
Round 3	528	0.349***	0.370***	0.344***	0.206***
		(0.042)	(0.046)	(0.045)	(0.040)
Aggregate Index of Sales and Profits					
Round 1	476	0.324***	0.233**	0.331***	0.08
		(0.092)	(0.109)	(0.099)	(0.070)
Round 2	570	0.333***	0.287***	0.330***	0.237***
		(0.057)	(0.070)	(0.068)	(0.060)
Round 3	535	0.397***	0.382***	0.369***	0.213***
		(0.072)	(0.101)	(0.101)	(0.070)
Innovation Index					
Round 1	476	0.111***	0.109***	0.109***	0.105***
		(0.033)	(0.033)	(0.031)	(0.029)
Round 2	528	0.166***	0.151***	0.172***	0.126***
		(0.027)	(0.030)	(0.031)	(0.028)
Round 3	466	0.138***	0.171***	0.201***	0.141***
		(0.031)	(0.034)	(0.033)	(0.029)

Notes: Standard errors in parentheses, *, **, *** indicate significance at the 10, 5, and 1 percent levels respectively. Pre-specified propensity score matching uses kernel matching within the common support and bootstraps standard errors. Nearest neighbor matching uses the Abadie et al. (2004) matching approach.

Appendix 6: Regression Discontinuity Estimates of Impact of 4-day training alone

Selection for the 4-day business plan training course was done on the basis of the initial application score, with scoring cutoffs which varied by region and new or existing firm used to determine the 6,000 firms selected for training. I use this feature to conduct a regression-discontinuity analysis of the impact of the training alone on firm outcomes.

To do this, I restrict attention to applicants in three regions: the North-Central, South-Eastern, and South-Western regions, since the other regions had few firms close to the eligibility cutoffs. In total this gave 4,008 new enterprises and 652 existing enterprises that had scores within 5 points on either side of the cutoff for being selected for business plan training. 770 of these firms (329 existing, 441 new) are already included in the sample of winners and the experimental control group. This leaves up to 3890 firms that could have been added to the survey. Given budget constraints, I chose to add all 323 existing firms, and then a random sample of 500 of the new firms with scores around the threshold. As a further complication, some of those who just made the cut-off for the training program then went on to be selected as winners, receiving the large capital grants. I exclude these firms, assuming they are selected at random (which many of them were).

The response rate for new firms for this regression discontinuity sample is 69.9% in round 1, 84.5% in round 2, and 82.0% in round 3. For existing firms it was 72.3% in round 1, 83.8% in round 2, and 85.4% in round 3.

Appendix Figures A6.1 and A6.2 then show graphically how the likelihood of being selected for training varies at the score cutoffs, and how key outcomes also change at these cutoffs. We see that while not being completely sharp, there is a large jump in the likelihood of being selected for training at the scoring cutoff.[3] I therefore use a fuzzy-RD design. Visually one then sees little change in the outcomes of interest around this threshold.

My pre-specified approach was to use the sample of non-winners within 5 points on either side of the cut-off I use instrumental variables to estimate the following regression:

$$Outcome_i = a + b * InvitedtoTraining_i + c * Region_i + d * mark_i + \varepsilon_i \qquad \text{(A1)}$$

Where *InvitedtoTraining* is instrumented with being above the scoring threshold. Since I am only looking within a very narrow window of the score around the threshold, I estimate equation (A1) with and without a linear control in the initial application mark. I complement this using the local regression approach of Calonico et al. (2014), implemented in Stata with the rdrobust command.

[3] I believe the lack of a complete sharp break reflects dynamic use of the scoring threshold to select invitees in batches subject to the constraint of 6,000 applicants overall to be selected, but only have the average threshold used.

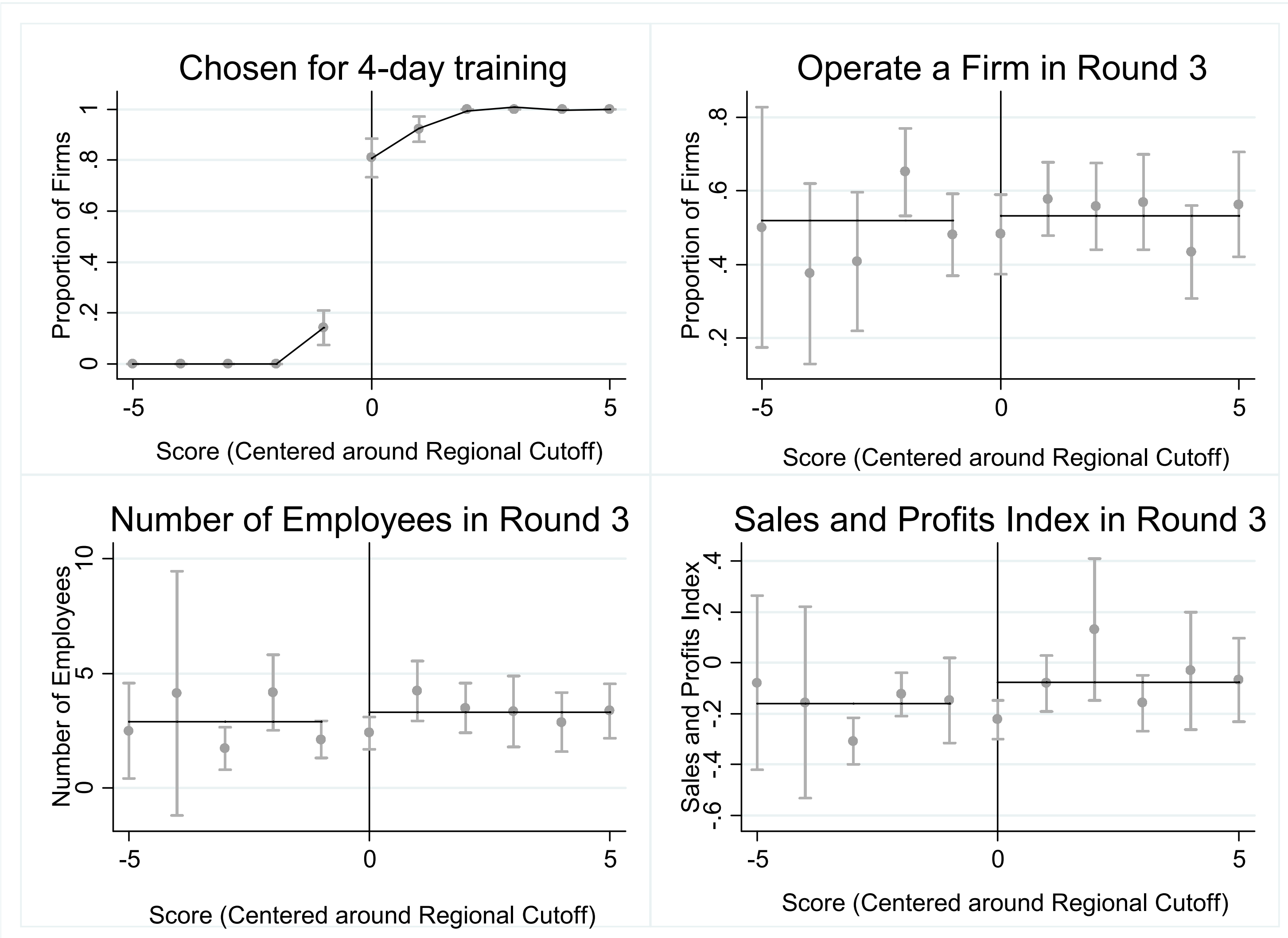
The results are shown in appendix tables A6.1 and A6.2 below. The first row shows having a score above the threshold is a strong and significant predictor of being invited to the training course, with this effect varying between 74 and 90 percentage points depending on the specification used. Table A6.1 then shows that there is no significant impact of the 4-day training on new firms for any of the five outcomes examined (operating a firm, total employment, having 10 or more workers, the sales and profits index, and the innovation index) in any of three survey rounds. Moreover, the point estimates are small in magnitude. The top of a 90 percent confidence interval in round 3 for employment is 1.59 workers and for 10+ workers is a 5.4 percentage point increase – which are both only approximately one quarter of the experimental treatment effects.

Table A6.2 shows there is also little in the way of significant impacts of the 4-day training on existing firms, although the estimates are more sensitive to functional form. In the round 3 data, we see a significant positive impact on the likelihood of operating a firm with a linear control or local regression. Looking at the graph in Figure A6.2, we see that exactly at the threshold more applicants have surviving firms, with this then dropping at one point past the threshold. If we use the "donut-hole" approach to RD to drop the observations right at the threshold, the estimated effect drops to 0.07 (p=0.37). Moreover, if anything, the firms which went through training have fewer workers in round 3. It therefore does not seem that the training program can account for the program impacts.

These findings are consistent with those in McKenzie and Woodruff (2014), who review RCTs of business training programs and find that many programs struggle to show effects. They suggest several reasons for this lack of significant impact. The first is that short programs of just a few days struggle to change enough in the way the business is operated to generate detectable impacts. The second, related, issue is one of statistical power – small samples and firm heterogeneity makes it difficult to detect impacts even when they do occur. This is more of an issue for the existing firm results here, which have a smaller sample and wider standard errors on the estimates than is the case for new firms. Nevertheless, even for existing firms, the confidence intervals do not contain very large impacts on employment or the likelihood of surpassing 10 or more employees.

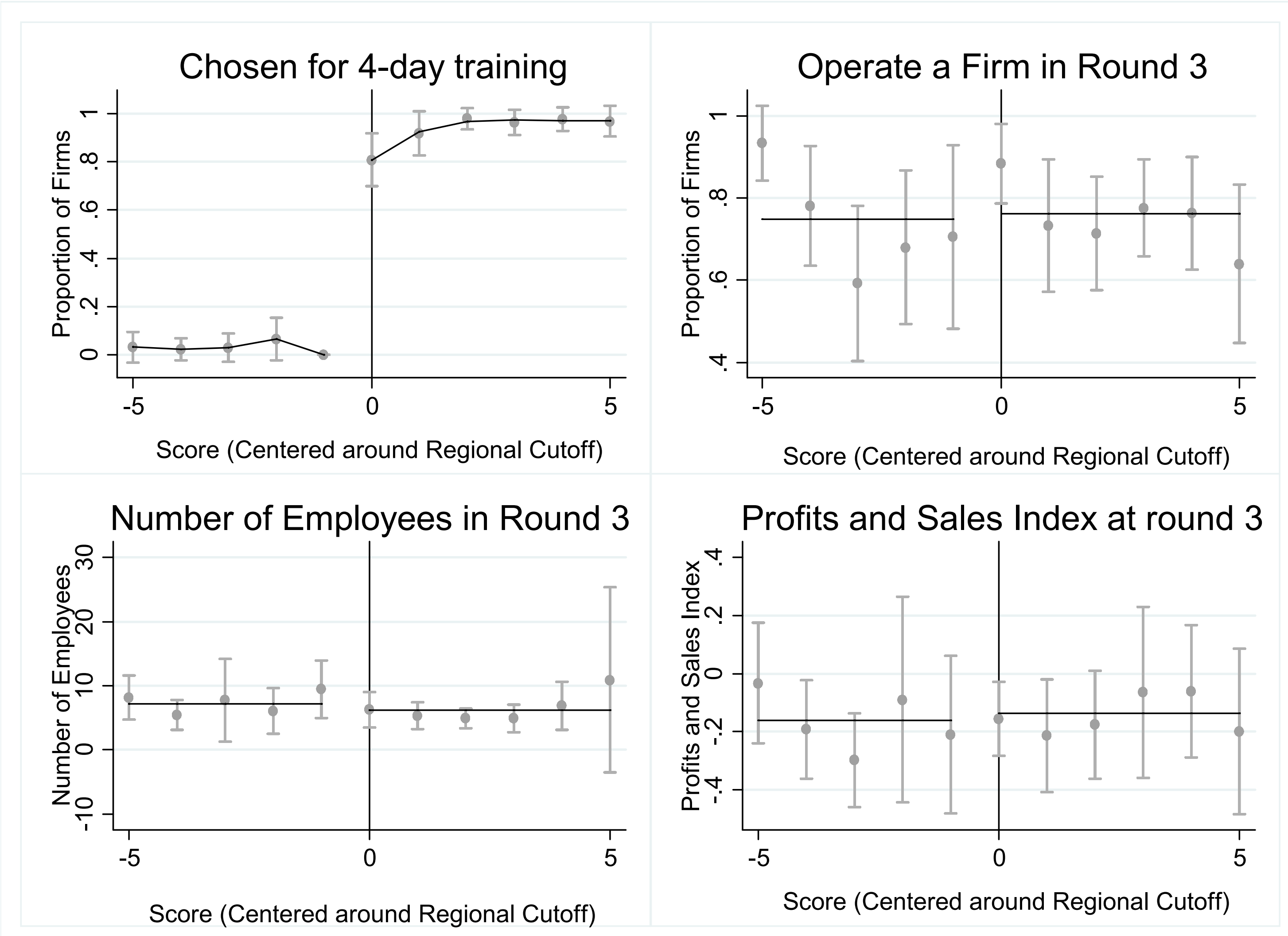
Finally, I note that I am unable to test directly for whether there is complementarity between the business plan training and the grant, since everyone who received the grant also obtained training. It is therefore possible that even though the training was not effective by itself, it may have enhanced the returns to the grant. If this were the case, I would expect to see improvements in business practices, which appear to be only small, suggesting this complementarity may not be that important in practice.

Appendix Figure A6.1: Regression Discontinuity Around the 4-Day Training Cutoff Impacts for New Firms



Notes: 95 percent confidence intervals shown around each score for mean at that point. Lines plotted are 4th order local polynomial for the chosen for 4-day training outcome, and local means on either side of the cutoff for the other outcomes.

Appendix Figure A6.2: Regression Discontinuity Around the 4-Day Training Cutoff Impacts for Existing Firms



Notes: 95 percent confidence intervals shown around each score for mean at that point. Lines plotted are 4th order local polynomial for the chosen for 4-day training outcome, and local means on either side of the cutoff for the other outcome

Appendix Table A6.1: Regression Discontinuity Estimates of Impact of 4-day business training on New Firms

	Sample Size	2SLS Regression Estimate	2SLS with Linear control	Calonico et al. (2014)	Experimental Estimate for Comparison
Invited to Training	772	0.881***	0.742***	0.829***	
		(0.019)	(0.038)	(0.025)	
Operates a firm at time of survey					
Round 1	509	0.064	0.089	0.055	0.213***
		(0.054)	(0.095)	(0.063)	(0.029)
Round 2	641	-0.013	0.039	-0.021	0.358***
		(0.049)	(0.086)	(0.056)	(0.023)
Round 3	628	0.002	0.040	0.007	0.373***
		(0.050)	(0.095)	(0.058)	(0.024)
Total Employment					
Round 1	493	0.450	0.412	0.276	1.426*
		(0.434)	(0.758)	(0.524)	(0.732)
Round 2	634	0.626	0.947	0.650	6.012***
		(0.455)	(0.861)	(0.540)	(0.412)
Round 3	594	0.342	1.135	0.648	5.227***
		(0.544)	(0.906)	(0.575)	(0.469)
Firm of 10+ Workers					
Round 1	493	0.018	0.023	0.013	0.024
		(0.024)	(0.042)	(0.028)	(0.020)
Round 2	634	0.027	0.036	0.029	0.288***
		(0.023)	(0.043)	(0.028)	(0.026)
Round 3	594	0.016	-0.018	0.002	0.229***
		(0.028)	(0.051)	(0.033)	(0.028)
Aggregate Index of Sales and Profits					
Round 1	500	0.077	0.031	0.045	0.016
		(0.062)	(0.114)	(0.079)	(0.047)
Round 2	629	0.043	0.060	0.030	0.298***
		(0.067)	(0.127)	(0.076)	(0.036)
Round 3	622	0.074	0.016	0.070	0.167***
		(0.067)	(0.115)	(0.074)	(0.042)
Innovation Index					
Round 1	500	0.047	-0.003	0.020	0.099***
		(0.030)	(0.053)	(0.035)	(0.019)
Round 2	606	0.027	0.010	0.010	0.270***
		(0.027)	(0.046)	(0.031)	(0.018)
Round 3	549	0.017	0.086	0.037	0.219***
		(0.030)	(0.060)	(0.036)	(0.019)

Notes: Standard errors in parentheses, *, **, *** indicate significance at the 10, 5, and 1 percent levels respectively. 2SLS regressions control for regional dummies, and differ only on whether or not a linear control in the initial application score is included as a control. Calonico et al. (2014) are with a triangular kernel, bandwidth of 5, using the rdrobust command in Stata with local mean regression.

Appendix Table A6.2: Regression Discontinuity Estimates of Impact of 4-day business training on Existing Firms

	Sample Size	2SLS Regression Estimate	2SLS with Linear control	Calonico et al. (2014)	Experimental Estimate for Comparison
Invited to Training	433	0.902***	0.772***	0.868***	
		(0.020)	(0.059)	(0.029)	
Operates a firm at time of survey					
Round 1	305	-0.004	-0.007	0.008	0.082***
		(0.039)	(0.100)	(0.051)	(0.027)
Round 2	281	0.025	0.090	0.035	0.130***
		(0.063)	(0.100)	(0.070)	(0.025)
Round 3	294	0.111	0.230**	0.135*	0.196***
		(0.076)	(0.114)	(0.082)	(0.031)
Total Employment					
Round 1	293	-1.085	-2.791	-2.351*	1.461*
		(1.053)	(2.591)	(1.424)	(0.808)
Round 2	276	-4.560	-2.531	-3.812	2.521*
		(3.000)	(3.402)	(2.837)	(1.366)
Round 3	267	-1.613	-4.978	-2.816	4.391***
		(2.286)	(3.424)	(1.934)	(0.674)
Firm of 10+ Workers					
Round 1	293	-0.044	-0.152	-0.121*	0.055
		(0.054)	(0.136)	(0.070)	(0.041)
Round 2	276	-0.092	-0.046	-0.100	0.211***
		(0.073)	(0.124)	(0.084)	(0.041)
Round 3	267	-0.113	-0.177	-0.161*	0.206***
		(0.074)	(0.120)	(0.083)	(0.040)
Aggregate Index of Sales and Profits					
Round 1	296	0.006	0.003	-0.029	0.08
		(0.096)	(0.247)	(0.141)	(0.070)
Round 2	277	-0.055	0.088	-0.009	0.237***
		(0.136)	(0.197)	(0.141)	(0.060)
Round 3	290	0.073	-0.132	0.021	0.213***
		(0.119)	(0.186)	(0.123)	(0.070)
Innovation Index					
Round 1	296	0.004	-0.053	-0.026	0.105***
		(0.039)	(0.096)	(0.050)	(0.029)
Round 2	261	0.035	0.033	0.031	0.126***
		(0.045)	(0.084)	(0.052)	(0.028)
Round 3	251	0.027	0.124	0.058	0.141***
		(0.053)	(0.088)	(0.057)	(0.029)

Notes: Standard errors in parentheses, *, **, *** indicate significance at the 10, 5, and 1 percent levels respectively. 2SLS regressions control for regional dummies, and differ only on whether or not a linear control in the initial application score is included as a control. Calonico et al. (2014) are with a triangular kernel, bandwidth of 5, using the rdrobust command in Stata with local mean regression.

Appendix 7: Further Details on Survey Methodology and Robustness to Survey Type

The main form of survey was an in-person interview conducted in English. English is an official language of Nigeria, and since the applicants had all submitted their applications in English and had high levels of education, English was used for all surveys.

Trained enumerators visited the applicants. During the first two survey rounds they contacted the applicants and interviewed them at either their business, their home, or a third location such as at the survey team's headquarters. Those who were not operating businesses would of course not be able to be interviewed at the business, but even those operating businesses sometimes preferred to be interviewed at home (where they would not be talking in front of employees, would have more time to focus on the survey, and would not be interrupted by business activities) or at the survey team's headquarters (this helped alleviate any trust issues about whether this was a genuine survey). For the third round survey, once trust had been built up, we emphasized much more strongly interviewing at the business premises. This increased the share of in-person interviews taking place at the business location (of those operating a business) from 48-49 percent in rounds 1 and 2 to 90 percent in round 3.

For those applicants who could not be interviewed using the in-person survey, two approaches were used. The survey enumerator attempted to directly observe whether a business was in operation and how many employees there were, and also used proxy respondents of neighboring business owners and family members to verify this information. 4.3 percent of the data on business operation in round 1 was obtained this way, as was 1.6 percent in round 2 and 1.1 percent in round 3. Secondly, in rounds 2 and 3, a shorter phone and online survey was offered to those who refused the longer in-person survey or were in areas where security restrictions made it difficult to interview. This shorter survey collected data on whether the business was operating, employment, sales and profits, and employment status and wage earnings of those not operating firms. 11.5 percent of the data in round 2 and 17.7 percent of the data in round 3 were collected with this shorter survey, allowing the overall response rates to be higher in these later two rounds than in the first round. In the long-term follow-up, 15.1 percent of the data comes from the shorter survey, and an additional 12.8% of observations on firm operation come from refusals where only whether or not a business was operating could be ascertained.

The method of interview is endogenous to business outcomes, since those not operating a business could not be interviewed at their place of business, and factors such as how busy the business was, how much they trust outsiders, etc. may have jointly determined whether or not they only were willing to take the shorter survey instrument as well as business outcomes. I therefore do not control for survey method in my main regressions, since this would involve controlling for a potentially endogenous variable. Nevertheless, in practice the results are robust to this choice: Appendix Table A7 shows the impacts on business operation, employment, the innovation index, and the sales and profits index are robust to controlling for whether or not the data on operation was obtained by proxy, and whether or not it was obtained by the short phone and online survey.

Appendix Table A7: Robustness of Main Results to Controlling for Survey Type

	New Firms				Existing Firms			
	Operate Firm	Total Employment	Innovation Index	Profits and Sales Index	Operate Firm	Total Employment	Innovation Index	Profits and Sales Index
Round 1								
Without survey type controls	0.215***	1.426*	0.099***	0.016	0.083***	1.512*	0.105***	0.080
	(0.029)	(0.732)	(0.019)	(0.047)	(0.027)	(0.795)	(0.029)	(0.070)
With survey type controls	0.216***	1.353*	0.096***	0.008	0.077***	1.361*	0.096***	0.056
	(0.029)	(0.733)	(0.019)	(0.047)	(0.027)	(0.793)	(0.029)	(0.070)
Round 2								
Without survey type controls	0.359***	6.012***	0.270***	0.298***	0.130***	2.556*	0.126***	0.237***
	(0.023)	(0.412)	(0.018)	(0.036)	(0.025)	(1.388)	(0.028)	(0.060)
With survey type controls	0.354***	6.019***	0.251***	0.293***	0.115***	2.336	0.110***	0.222***
	(0.023)	(0.406)	(0.018)	(0.036)	(0.024)	(1.470)	(0.028)	(0.060)
Round 3								
Without survey type controls	0.373***	5.227***	0.219***	0.167***	0.195***	4.425***	0.141***	0.211***
	(0.024)	(0.469)	(0.019)	(0.042)	(0.031)	(0.673)	(0.029)	(0.070)
With survey type controls	0.375***	5.226***	0.206***	0.166***	0.193***	4.377***	0.126***	0.206***
	(0.024)	(0.467)	(0.019)	(0.042)	(0.032)	(0.676)	(0.029)	(0.071)

Notes: survey type controls are a control for data on business operation and employment being reported by proxy, and for data being collected via a shorter survey instrument administered via phone or online.

Appendix 8: Robustness to Attrition

Appendix Table A8.1 reports the response rates by treatment status and round. I report this for four types of data: having information on whether or not a firm is in operation (which can be verified and collected from proxy respondents even when no survey is available), responding to the survey, and providing information on employment and profits.

Appendix Table A8.1: Response rates by treatment status and round

		All Firms			New Firms			Existing Firms		
	Sample Size	First Survey	Second Survey	Third Survey	First Survey	Second Survey	Third Survey	First Survey	Second Survey	Third Survey
Panel A: Information available on whether or not they operate a firm										
Treatment Group	729	0.846	0.930	0.885	0.840	0.922	0.869	0.856	0.942	0.910
Control Group	1112	0.752	0.906	0.825	0.756	0.901	0.816	0.738	0.924	0.852
Experimental Sample	1841	0.789	0.916	0.848	0.785	0.908	0.835	0.799	0.933	0.882
Panel B: Responded to the Survey										
Treatment Group	729	0.812	0.915	0.870	0.796	0.900	0.847	0.838	0.939	0.906
Control Group	1112	0.719	0.872	0.805	0.723	0.863	0.800	0.703	0.901	0.821
Experimental Sample	1841	0.756	0.889	0.831	0.748	0.876	0.816	0.773	0.921	0.865
Panel C: Data on Employment										
Treatment Group	729	0.812	0.925	0.868	0.796	0.914	0.851	0.838	0.942	0.896
Control Group	1112	0.735	0.886	0.784	0.740	0.880	0.777	0.719	0.905	0.806
Experimental Sample	1841	0.765	0.901	0.817	0.759	0.892	0.803	0.780	0.924	0.852
Panel D: Data on Profits										
Treatment Group	729	0.819	0.914	0.867	0.807	0.902	0.847	0.838	0.932	0.899
Control Group	1112	0.738	0.882	0.809	0.743	0.875	0.802	0.722	0.905	0.833
Experimental Sample	1841	0.770	0.895	0.832	0.765	0.885	0.818	0.782	0.919	0.867

I examine robustness to attrition in several ways. The first check is to determine whether the sample which responds to the surveys is balanced on observed baseline characteristics, and whether the characteristics of those who attrit differ between treatment and control. I do this for the last survey round. Appendix Tables A8.2 and A8.3 show that the treatment and control group respondents are well balanced on observable baseline characteristics for both new (p-value for orthogonality test 0.902) and existing (p-value for orthogonality test 0.956) firms. Moreover, in both cases we cannot reject that the characteristics of those who attrit are also jointly orthogonal to treatment status. These checks suggest that attrition is not causing the treatment and control groups to differ on pre-existing characteristics.

Appendix Table A8.2: Balance on Baseline Covariates by Response Status in Round 3 for New Firm Applicants

	Responded to Round 3			Attrited from Round 3		
	Treatment	Control	p-value	Treatment	Control	p-value
Applicant Characteristics						
Female	0.16	0.16	s	0.25	0.24	s
Age	29.4	29.5	0.647	29.0	29.9	0.238
Married	0.35	0.35	0.758	0.28	0.40	0.032
High School or Lower	0.10	0.09	0.958	0.17	0.14	0.792
University education	0.71	0.70	0.321	0.59	0.76	0.053
Postgraduate education	0.04	0.06	0.530	0.09	0.06	0.494
Lived Abroad	0.06	0.07	0.778	0.07	0.16	0.294
Choose Risky Option	0.56	0.54	0.358	0.59	0.59	0.437
Have Internet access at home	0.46	0.46	0.381	0.51	0.54	0.897
Own a Computer	0.84	0.86	0.635	0.86	0.85	0.419
Satelite Dish at home	0.69	0.65	0.221	0.62	0.62	0.757
Freezer at home	0.52	0.55	0.644	0.46	0.58	0.315
Business Characteristics						
Crop and Animal Sector	0.23	0.23	0.603	0.14	0.18	0.679
Manufacturing Sector	0.29	0.24	0.198	0.25	0.21	0.629
Trade Sector	0.03	0.04	0.624	0.06	0.06	0.819
IT Sector	0.07	0.06	0.533	0.06	0.07	0.685
First Round Application Score	59.9	60.1	0.173	59.5	59.1	0.772
Business Plan Score	53.6	55.2	0.731	54.0	56.4	0.939
Sample Size	382	679		69	170	
Joint orthogonality test: treatment vs control			0.902			0.459

Notes: s denotes variable used in forming randomization strata. P-values are for test of equality of means after controlling for randomization strata.

Appendix Table A8.3: Balance on Baseline Covariates by Response Status in Round 3 for Existing Firm Applicants

	Responded to Round 3			Attrited from Round 3		
	Treatment	Control	p-value	Treatment	Control	p-value
Applicant Characteristics						
Female	0.17	0.16	s	0.27	0.23	s
Age	32.1	31.9	0.700	31.3	31.5	0.238
Married	0.51	0.55	0.387	0.42	0.60	0.032
High School or Lower	0.13	0.12	0.537	0.12	0.13	0.792
University education	0.62	0.66	0.398	0.77	0.70	0.053
Postgraduate education	0.07	0.11	0.115	0.19	0.15	0.494
Lived Abroad	0.09	0.10	0.668	0.19	0.17	0.294
Choose Risky Option	0.56	0.51	0.334	0.65	0.57	0.437
Have Internet access at home	0.55	0.61	0.236	0.73	0.64	0.897
Own a Computer	0.88	0.89	0.858	0.77	0.83	0.419
Satelite Dish at home	0.66	0.72	0.127	0.73	0.64	0.757
Freezer at home	0.56	0.61	0.308	0.65	0.62	0.315
Business Characteristics						
Crop and Animal Sector	0.16	0.14	0.540	0.12	0.21	0.679
Manufacturing Sector	0.28	0.28	0.904	0.23	0.15	0.629
Trade Sector	0.05	0.05	0.970	0.12	0.04	0.819
IT Sector	0.15	0.15	0.959	0.12	0.11	0.685
First Round Application Score	57.1	56.3	0.223	57.9	57.9	0.772
Business Plan Score	45.7	45.2	0.394	47.0	46.3	0.939
Number of Workers	7.03	7.65	0.428	10.04	8.10	0.852
Ever had Formal Loan	0.06	0.08	0.309	0.08	0.11	0.903
Sample Size	252	216		26	47	
Joint orthogonality test: treatment vs control			0.956			0.995

Notes: s denotes variable used in forming randomization strata. P-values are for test of equality of means after controlling for randomization strata.

Although these results suggest that the treatment and control groups remain comparable, I examine sensitivity to attrition through a number of checks.

1. Lee (2009) bounds. This approach makes a monotonicity assumption and then adjusts for differential attrition between treatment and control. Since the response rates are higher for the treated group than the control group, the assumption here is that there are individuals who would attrit if they end up in the control group but not if they end up in the treatment group, but not vice versa. This appears plausible in our case if we think that part of attrition is not being able to be found because a business is not in operation (and treatment helps makes it more likely the business is in operation), and that people who have received money for their business may be happier to talk about this business. Then, for example, for new enterprises, the response rate in round 1 is 84

percent for treated firms, and 76 percent for control firms. The difference in response rates is thus 8 percent. I therefore trim 8/84 = 10 percent of the treated observations, with the lower bound occurring when I trim observations which are operating a business, and the upper bound when I trim observations which are not operating a business. In the case of existing firms, there are insufficient closed firms to cover the number required to be trimmed, so I also choose a random sample of the open firms to trim. These bounds are widest for round 1, and narrower for subsequent rounds due to higher survey response rates and less differences in response rates by treatment status in future rounds.

2. Imbens and Manski (2004) confidence intervals. At the suggestion of a referee, I include these confidence intervals for the Lee bounds, which take account of sampling variability as well as the potential selection bias from differential attrition.

3. Behagel et al. (2015) bounds. The idea here is to record the number of attempts made to interview firms, and then narrow the Lee bounds by trimming the harder to interview individuals from the treatment group. For example, in round 1, for new firms, 75.5 percent of the control group responded to the survey. Among the treatment group, 76.5 percent responded in four or fewer attempts. Lee bounds are then applied to this sample consisting of the first 75 to 76 percent of firms to be interviewed in each group. This approach does not narrow the bounds in rounds 2 and 3, since the response rates are higher for each group and the response rate for the control group was only reached on the last number of attempts for the treated group.

4. The impact controlling for the baseline application score and business plan score – this controls for any differential selective attrition on these two key observed assessments of business potential.

5. Horowitz and Manski (2000) worst case bounds. These bounds adjust for total attrition, rather than differential attrition. The worst case for a positive treatment effect is if all the missing treated firms are closed, and all the missing control firms are open. For the sales and profits index, the worst case has all missing control firms take on the maximum observed value of the index amongst treated firms. These bounds cover a situation that is unlikely to be relevant: for example, if we look at those existing firms who attrited in round 1, but were found in round 2, 94% of the treated and 83% of the control were operating businesses: so assuming that all the missing treated are closed is likely to be highly misleading.

6. Filling in missing data based on past closure status. This is done for round 3 only, and assumes that firms which were closed in round 1 and/or 2, and then not interviewed thereafter, remain closed.

These robustness checks show that the key results of the paper are robust to attrition, with only the worst case bounds overturning the results for some outcomes.

Table A8.4: Robustness of Impact on Start-up and Survival to Attrition

Robustness of Impact on Start-up to Alternative Ways of Dealing with Attrition

	New Firms			Existing Firms		
	Round 1	Round 2	Round 3	Round 1	Round 2	Round 3
Table 2 Impact for Comparison	**0.215*** **	**0.359*** **	**0.373*** **	**0.083*** **	**0.130*** **	**0.195*** **
	(0.029)	**(0.023)**	**(0.024)**	**(0.027)**	**(0.025)**	**(0.031)**
Lee (2009) bounds						
Upper bound	0.189***	0.356***	0.366***	0.074***	0.129***	0.191***
	(0.031)	(0.023)	(0.025)	(0.028)	(0.025)	(0.032)
Lower bounds	0.295***	0.379***	0.431***	0.127***	0.149***	0.238***
	(0.028)	(0.022)	(0.021)	(0.024)	(0.024)	(0.029)
Imbens and Manski (2004) Confidence Intervals	[0.133, 0.362]	[0.312, 0.422]	[0.328, 0.489]	[0.028, 0.169]	[0.085, 0.209]	[0.142, 0.288]
Behagel et al. (2012) bounds						
Upper bound	0.232***			0.071**		
	(0.030)			(0.028)		
Lower bound	0.217***			0.122***		
	(0.030)			(0.023)		
Impact controlling for business plan and application scores	0.213***	0.358***	0.373***	0.087***	0.134***	0.199***
	(0.029)	(0.023)	(0.024)	(0.028)	(0.026)	(0.032)
Horowitz and Manski (2000) worst case bounds	-0.000	0.248***	0.173***	-0.088***	0.061**	0.076**
	(0.028)	(0.024)	(0.026)	(0.029)	(0.027)	(0.032)
Filling in based on past closed status			0.390***			0.211***
			(0.033)			(0.023)

Notes: Imbens and Manski (2004) confidence intervals produced by Stata's leebounds command, and capture both uncertainty about the selection bias and the sampling error. Behagel et al. (2012) bounds use the number of attempts made to reach successful respondents to form bounds. This only sharpens the bounds compared to Lee (2009) bounds in round 1, since response rates overlap for the last attempts made in rounds 2 and 3. Impact controlling for business plan and application scores adds these baseline variables as controls in the treatment regression. Horowitz and Manski (2000) worst case bounds assume all missing control observations are from open businesses and all missing treated observations are from closed businesses. Filling in based on past closed status assumes firms closed in round 1 or 2 that are not interviewed in subsequent rounds remain closed in round 3.

Table A8.5: Robustness of Impact on Having 10+ Employees to Attrition

	New Firms			Existing Firms		
	Round 1	Round 2	Round 3	Round 1	Round 2	Round 3
Table 3 Impact for Comparison	**0.024**	**0.288*** **	**0.229*** **	**0.057**	**0.215*** **	**0.208*** **
	(0.020)	**(0.026)**	**(0.028)**	**(0.041)**	**(0.041)**	**(0.040)**
Lee (2009) bounds						
Upper bound	0.033	0.303***	0.261***	0.102**	0.232***	0.252***
	(0.021)	(0.027)	(0.029)	(0.044)	(0.042)	(0.042)
Lower bounds	-0.046***	0.266***	0.167***	-0.066*	0.192***	0.138***
	(0.015)	(0.026)	(0.028)	(0.038)	(0.041)	(0.040)
Imbens and Manski (2004) Confidence Intervals	[-0.102, 0.069]	[0.220, 0.354]	[0.116, 0.320]	[-0.172, 0.171]	[0.117, 0.308]	[0.052, 0.320]
Behagel et al. (2012) bounds						
Upper bound	0.026			0.053		
	(0.021)			(0.042)		
Lower bound	0.015			0.013		
	(0.020)			(0.041)		
Impact controlling for business plan and	0.024	0.288***	0.229***	0.058	0.217***	0.210***
application scores	(0.020)	(0.026)	(0.028)	(0.041)	(0.041)	(0.040)
Horowitz and Manski (2000) worst case bounds	-0.211***	0.152***	-0.013	-0.212***	0.120***	0.007
	(0.021)	(0.027)	(0.027)	(0.039)	(0.041)	(0.041)
Filling in based on past closed status			0.229***			0.209***
			(0.027)			(0.039)

Notes: Imbens and Manski (2004) confidence intervals produced by Stata's leebounds command, and capture both uncertainty about the selection bias and the sampling error. Behagel et al. (2012) bounds use the number of attempts made to reach successful respondents to form bounds. This only sharpens the bounds compared to Lee (2009) bounds in round 1, since response rates overlap for the last attempts made in rounds 2 and 3. Impact controlling for business plan and application scores adds these baseline variables as controls in the treatment regression. Horowitz and Manski (2000) worst case bounds assume all missing control observations are from businesses with 10+ workers and all missing treated observations are from firms with less than 10+ workers. Filling in based on past closed status assumes firms closed in round 1 or 2 that are not interviewed in subsequent rounds remain closed in round 3 and so have fewer than 10 workers then.

Table A8.6: Robustness of Impact on Innovation Index to Attrition

	New Firms			Existing Firms		
	Round 1	Round 2	Round 3	Round 1	Round 2	Round 3
Table 3 Impact for Comparison	**0.099*** **	**0.270*** **	**0.219*** **	**0.105*** **	**0.126*** **	**0.141*** **
	(0.019)	**(0.018)**	**(0.019)**	**(0.029)**	**(0.028)**	**(0.029)**
Lee (2009) bounds						
Upper bound	0.053***	0.266***	0.217***	0.047*	0.119***	0.114***
	(0.018)	(0.018)	(0.019)	(0.028)	(0.028)	(0.029)
Lower bounds	0.124***	0.284***	0.227***	0.174***	0.135***	0.171***
	(0.019)	(0.018)	(0.019)	(0.028)	(0.028)	(0.029)
Imbens and Manski (2004) Confidence Intervals	[0.000, 0.158]	[0.225, 0.319]	[0.141, 0.248]	[-0.015, 0.246]	[0.064, 0.199]	[0.051, 0.238]
Behagel et al. (2012) bounds						
Upper bound	0.110***			0.113***		
	(0.020)			(0.029)		
Lower bound	0.092***			0.056*		
	(0.019)			(0.029)		
Impact controlling for business plan and application scores	0.098***	0.269***	0.219***	0.107***	0.127***	0.142***
	(0.019)	(0.018)	(0.019)	(0.029)	(0.028)	(0.029)
Horowitz and Manski (2000) worst case bounds	-0.140***	0.054***	-0.125***	-0.144***	-0.044	-0.142***
	(0.020)	(0.020)	(0.021)	(0.030)	(0.029)	(0.031)
Filling in based on past closed status			0.218***			0.152***
			(0.018)			(0.029)

Notes: Imbens and Manski (2004) confidence intervals produced by Stata's leebounds command, and capture both uncertainty about the selection bias and the sampling error. Behagel et al. (2012) bounds use the number of attempts made to reach successful respondents to form bounds. This only sharpens the bounds compared to Lee (2009) bounds in round 1, since response rates overlap for the last attempts made in rounds 2 and 3. Impact controlling for business plan and application scores adds these baseline variables as controls in the treatment regression. Horowitz and Manski (2000) worst case bounds assume all missing control observations have 100% innovation and all missing treated observations are from firms with 0 innovation. Filling in based on past closed status assumes firms closed in round 1 or 2 that are not interviewed in subsequent rounds remain closed in round 3 and so have 0 innovation.

Table A8.7: Robustness of Impact on Sales and Profits Index to Attrition

	New Firms			Existing Firms		
	Round 1	Round 2	Round 3	Round 1	Round 2	Round 3
Table 4 Impact for Comparison	**0.016**	**0.298*****	**0.167*****	**0.080**	**0.237*****	**0.211*****
	(0.047)	**(0.036)**	**(0.042)**	**(0.070)**	**(0.060)**	**(0.070)**
Lee (2009) bounds						
Upper bound	0.059	0.328***	0.200***	0.204***	0.268***	0.281***
	(0.048)	(0.036)	(0.043)	(0.070)	(0.060)	(0.071)
Lower bounds	-0.114***	0.226***	0.080**	-0.099	0.160***	0.051
	(0.035)	(0.031)	(0.037)	(0.062)	(0.054)	(0.052)
Imbens and Manski (2004) Confidence Intervals	[-0.190, 0.145]	[0.148, 0.383]	[0.011, 0.273]	[-0.219, 0.339]	[0.043, 0.374]	[-0.055, 0.404]
Behagel et al. (2012) bounds						
Upper bound	0.038			0.125*		
	(0.048)			(0.070)		
Lower bound	-0.041			-0.040		
	(0.038)			(0.065)		
Impact controlling for business plan and application scores	0.016	0.296***	0.165***	0.082	0.240***	0.211***
	(0.047)	(0.036)	(0.043)	(0.071)	(0.060)	(0.070)
Horowitz and Manski (2000) worst case bounds	-1.631***	-1.341***	-1.199***	-1.398***	-0.207**	-1.244***
	(0.107)	(0.154)	(0.099)	(0.142)	(0.090)	(0.195)
Filling in based on past closed status			0.180***			0.230***
			(0.041)			(0.070)

Notes: Imbens and Manski (2004) confidence intervals produced by Stata's leebounds command, and capture both uncertainty about the selection bias and the sampling error. Behagel et al. (2012) bounds use the number of attempts made to reach successful respondents to form bounds. This only sharpens the bounds compared to Lee (2009) bounds in round 1, since response rates overlap for the last attempts made in rounds 2 and 3. Impact controlling for business plan and application scores adds these baseline variables as controls in the treatment regression. Horowitz and Manski (2000) worst case bounds assume all missing control observations have the highest observed level of the index and all missing treated are from firms with the lowest level. Filling in based on past closed status assumes firms closed in round 1 or 2 that are not interviewed in subsequent rounds remain closed in round 3 and so have no profits or sales.

Appendix 9: Measurement of Key Outcomes

The survey instruments and data are available in the World Bank's open data library (the replication files will be added once a final version of the paper is accepted):

http://microdata.worldbank.org/index.php/catalog/2329

Nominal Naira were converted into real (November 2012) Naira using the Consumer CPI of the Central Bank of Nigeria.

Key outcomes are measured as follows:

Operates a firm: measuring by directly asking the owner if they currently operate a business. When the owner could not be found, this information was obtained by direct verification and through asking friends, relatives, and neighbors.

Own employment: a binary variable taking the value one if the owner owns a business, or reports working in the last month for pay in any other occupation.

Total employment: the number of paid workers in the firm, including the owner. This is the sum of whether the owner operates a firm, the number of wage and salary employees, and the number of casual and daily laborers. Unpaid workers are not included. Coded as zero if the business does not exist.

Firm of 10+ workers: a binary variable taking the value one if total employment is 10 or more.

Truncated sales: total sales reported in the last month, truncated at the 99th percentile to reduce the influence of outliers.

Truncated profits: total profits of the business in the last month, truncated at the 99th percentile. Profits are measured following the recommendations of de Mel et al. (2009) by means of a direct question asking "What was the total income the business earned during the last month after paying all expenses including wages of employees, but not including any income you paid yourself". They therefore include the return to the entrepreneur's own labor.

Inverse hyperbolic sine of profits: the transform $\log(y+(y^2+1)^{1/2})$ which is similar to the log transformation but which allows for zero and negative values of profits.

Aggregate Index of Sales and Profits: an average of standardized z-scores for the outcomes monthly sales, truncated monthly sales, annual sales, sales higher than one year ago, monthly profits, truncated monthly profits, profits in the best month, and inverse hyperbolic sine of profits.

Innovation Index: an average of standardized z-scores for the following 12 different measures of innovative activities:

- Introduced a new product in the past year
- Improved an existing product or service in the past year
- Introduced a new or improved process in the past year
- Introduced a new design or packaging in the past year
- Introduced a new channel for selling goods in the past year
- Introduced a new method of pricing in the past year
- Introduced a new method of advertising in the past year
- Changed the way work is organized in the firm in the past year
- Introduced new quality control standards in the past year
- Licensed a new technology in the past year
- Obtained new quality certification in the past year
- Uses the internet

Took a formal loan: received a loan from a bank, microfinance organization, or NGO in the past year

Received equity investment: received a new investment in the form of equity in the past year

Value of inventories: current value reported of inventories and raw materials, top-coded at the 99^{th} percentile.

Made a large capital purchase: reports making a capital purchase of more than 100,000 Naira in the past year

Value of capital purchases: total value of capital purchases over 100,000 Naira, truncated at the 99^{th} percentile.

Capital stock: current value of inventories plus the sum of the value of capital purchases made in each of the survey rounds since the program began, truncated at the 99^{th} percentile.

Entrepreneurial Self-Efficacy: Measured as the number of 9 business activities that the owner rates themselves as "very confident" in their ability to do. This is coded as 1 for each item if the owner answers 4 = very confident, and 0 if they answer 1 through 3, or 9 (not applicable or refuse). The activities are:

- Come up with an idea for a new business product or service
- Estimate accurately the costs of a new business venture
- Estimate customer demand for a new product or service
- Sell a product or service to a customer you are meeting for the first time
- Identify good employees who can help the business grow
- Inspire, encourage, and motivate employees
- Find suppliers who will sell you raw materials at the best price
- Persuade a bank to lend you money to finance a business venture
- Correctly value a business if you were to buy an existing business from someone else.

Has a Mentor: The firm reports have a business mentor in response to a direct question.

Number in Business Network: number of other business owners the individual discusses business matters with, truncated at the 99^{th} percentile.

Firm is formal: the firm reports that it has a registered business name, a municipal license, and pays income tax

Hours of consulting services: number of hours of consulting services used in the past year, truncated at the 99^{th} percentile.

Functioning website: has a website with a functioning URL as of October 4, 2015.

Business practices: the proportion of the following 22 business practices employed, taken from McKenzie and Woodruff (2015):

Marketing Practices: coded as 1 for each of the following that the business has done in the last 3 months:

- M1: Visited at least one of its competitor's businesses to see what prices its competitors are charging
- M2: Visited at least one of its competitor's businesses to see what products its competitors have available for sale
- M3: Asked existing customers whether there are any other products the customers would like the business to sell or produce
- M4: Talked with at least one former customer to find out why former customers have stopped buying from this business
- M5: Asked a supplier about which products are selling well in this business' industry
- M6: Attracted customers with a special offer
- M7: Advertised in any form (last 6 months)

Note: M1 and M2 are coded as zero if the firm says it has no competitors. M4 is coded as zero if the firm says it has no former customers.

Buying and Stock Control Practices: coded as 1 for each of the following:

- B1: Attempted to negotiate with a supplier for a lower price on raw material
- B2: Compared the prices or quality offered by alternate suppliers or sources of raw materials to the business' current suppliers or sources of raw material
- B3: The business does not run out of stock monthly or more (coded as one if the business has no stock)

Costing and Record-Keeping Practices: coded as 1 for each of the following that the business does:

- R1: Keeps written business records
- R2: Records every purchase and sale made by the business

- R3: Able to use records to see how much cash the business has on hand at any point in time
- R4: Uses records regularly to know whether sales of a particular product are increasing or decreasing from one month to another
- R5: Works out the cost to the business of each main product it sells
- R6: Knows which goods you make the most profit per item selling
- R7: Has a written budget, which states how much is owed each month for rent, electricity, equipment maintenance, transport, advertising, and other indirect costs to business
- R8: Has records documenting that there exists enough money each month after paying business expenses to repay a loan in the hypothetical situation that this business wants a bank loan

Financial Planning Practices: coded as 1 for each of the following:

- F1: Review the financial performance of their business and analyze where there are areas for improvement at least monthly
- F2: Has a target set for sales over the next year
- F3: Compares their sales achieved to their target at least monthly
- F4: Has a budget of the likely costs their business will have to face over the next year

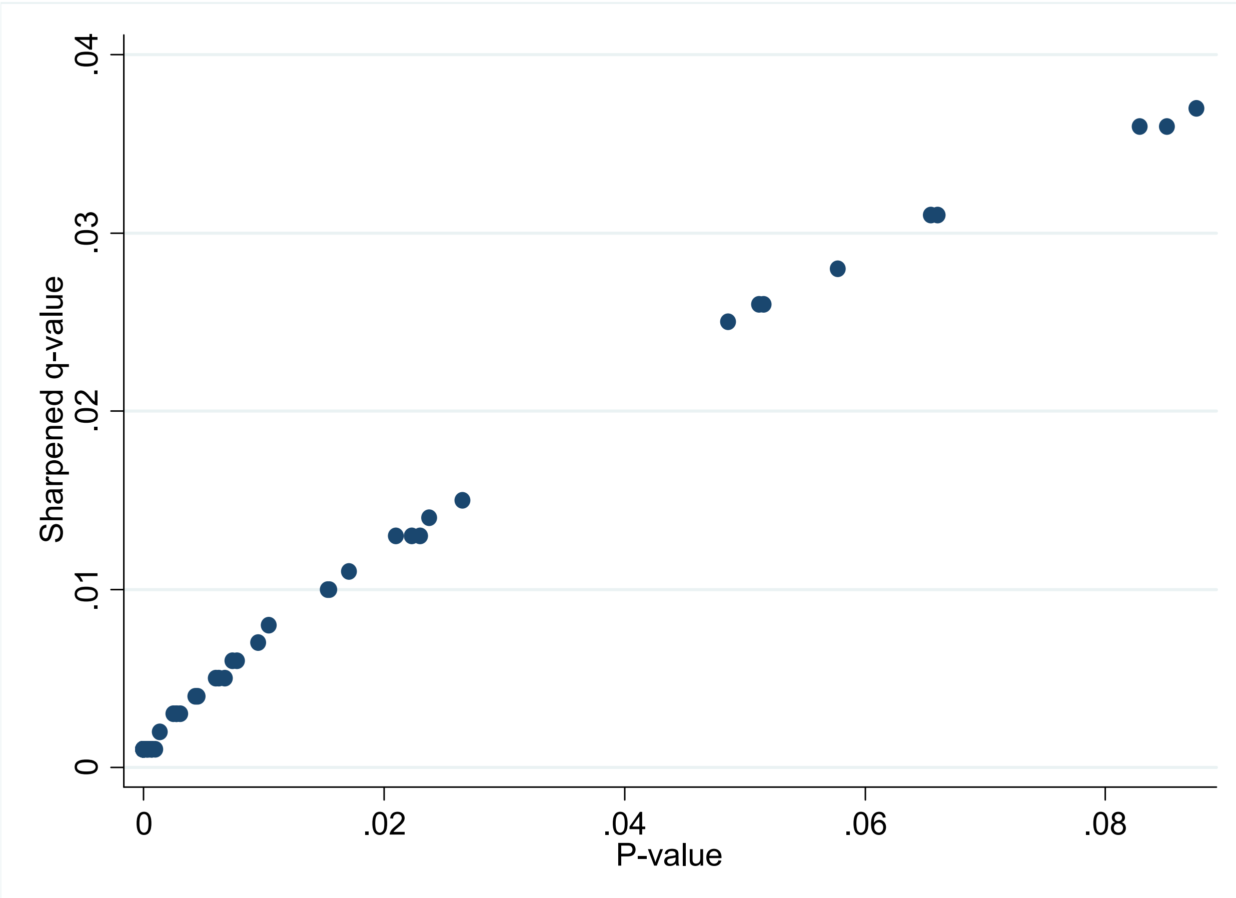
Appendix 10: Robustness to Corrections for Multiple Hypothesis Testing

The main approach I use to deal with multiple hypothesis testing is to group the outcomes into domains or families, and then examine the impact on an aggregate or standardized index measure within domain. For employment, total employment is a natural aggregate; for profits and sales I take an index of standardized z-scores; for capital, total capital stock is a natural aggregate; and I also use an index of innovative practices.

A second approach that can be used is to construct sharpened q-values following Anderson (2008) and Benjamini et al. (2006). This process uses a two-stage procedure to control the false discovery rate (the expected proportion of rejections that are type I errors) when reporting results for specific outcomes (e.g. for the outcome of 10+ workers for new firm applicants in the round 3 survey). I take the 130 p-values from the tests in Tables 2 through 8 and implement the procedure on this full set of p-values. Figure A10 below plots the resulting sharpened q-values against the p-values for all p-values that have unadjusted values of 0.100 or lower in the paper. We see that the all of these outcomes that are significant at the 10 percent level without adjustment are actually significant at the 5 percent level or better using the sharpened q-values. Note that the sharpened q-values are actually less than the p-values for p-values between 0.01 and 0.100 – the reason is that there are so many (82) hypotheses rejected with p-values below 0.01 that the false discovery rate can be controlled whilst still allowing for some false rejections in this range.[4] The key point to note is that all significant results reported in the paper are robust to using this adjustment for multiple hypothesis testing.

[4] Anderson (2008) notes how correctly rejecting some false hypotheses increases the power of the false discovery rate since this then lowers the expected proportion of all rejections which are type I errors. The specific point that sharpened q-values can be less than unadjusted p-values as a result appears in Michael Anderson's Stata code for sharpened q-values: http://are.berkeley.edu/~mlanderson/downloads/fdr_sharpened_qvalues.do.zip

Figure A10: Sharpened q-values vs p-values for Tables 2 8 with p-values<0.100



Appendix 11: Measurement of Employment, Impact on Job Creation, and Further Employment Results

Appendix 11.1: Reporting of Employment in Administrative Data vs the Survey

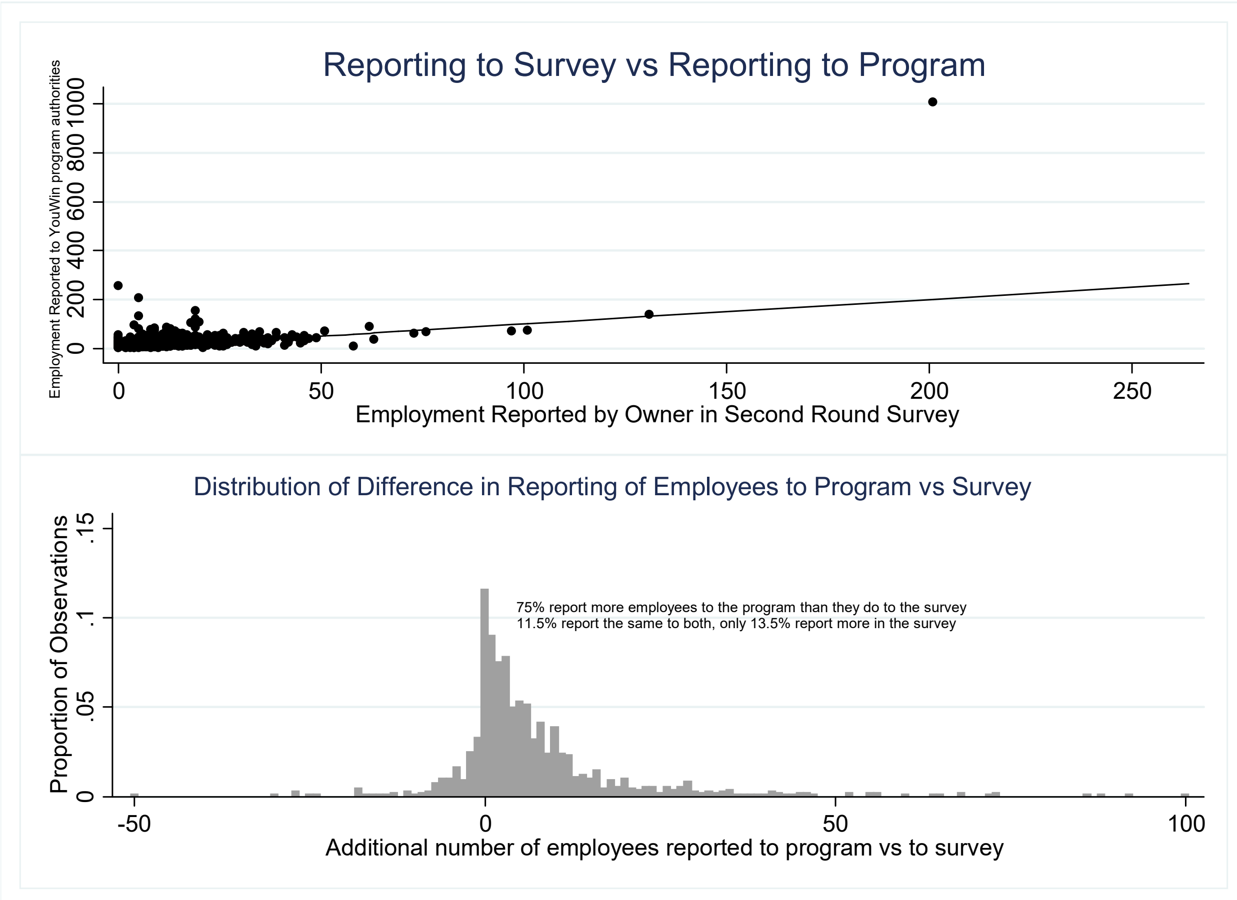
Since employment creation was a goal of the program, and firms were meant to show progress in growing their firms before receiving their last tranche payments, one may be concerned that firm owners are over-reporting employment. Around the time of the second round survey, firm owners were asked to report how many permanent employees, and how many casual or temporary employees they had employed since they won the grant. The total of 23,781 employees reported in the administrative data is almost twice the total employment of 13,945 reported to us in our surveys.

Appendix Figure A11.1 plots our survey measure of total employment against the total reported by firms to the program. We see many reports lie close to the 45 degree line, but there are numbers of firms with relatively few workers in the survey who report more to the program, as well as one firm reporting around 200 workers in the survey versus 1000 in the program. The bottom panel of this figure plots the distribution of differences: 75 percent of firms report more employees to the program than they do in the survey, 11.5 percent report the same to both, and only 13.5 percent report more in the survey than in the firm.

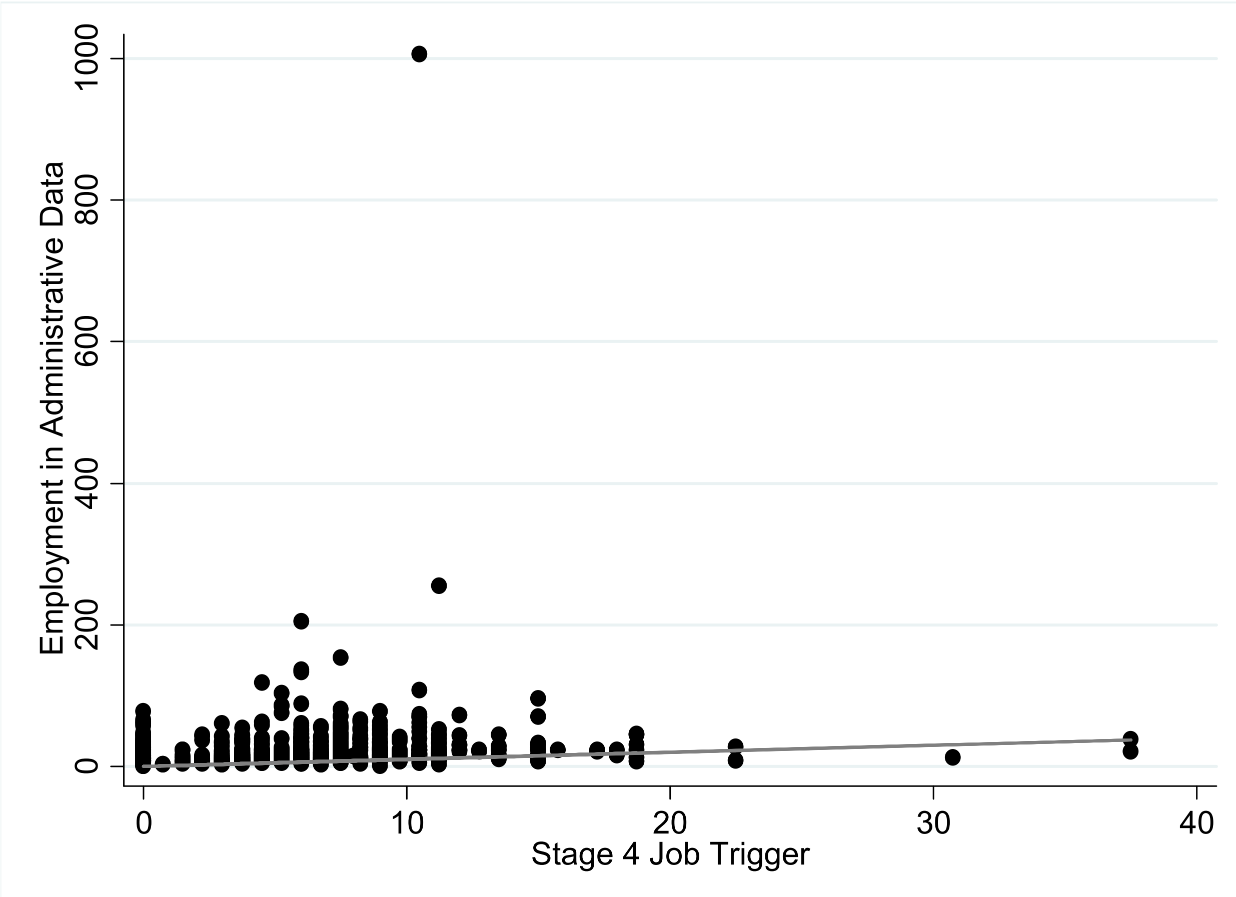
While firm owners may have an incentive to over-report employment to the program to ensure they reach the job triggers needed for their third and fourth tranche payments, these triggers were set very low and the amounts reported in the administrative data greatly exceed these triggers (appendix Figure A11.2), with the median firm having 9 more workers in the administrative data than needed for their fourth tranche payment to have been made.

Firm owners have an incentive to over-report employment to the program, whereas this incentive is much less in the survey which was conducted by a survey research firm (TNS Gallup) and where the questions came as part of a much more detailed set of questions about the business. I therefore view the survey measures as more reliable. As an added check on this, survey enumerators were asked to record how many employees they physically observed at the enterprise while they were conducting the interview. This misses workers who are sick, those whose hours don't correspond with those of the interview, and those who are working in another location. Furthermore, it is not available for individuals who were interviewed at their house, or over the phone. Nevertheless, it provides a useful check which I discuss in the text.

Appendix Figure A11.1: Comparing Administrative to Survey Reports on Employment



Appendix Figure A11.2: Employment Reported in Administrative Data Greatly Exceeds the Job Trigger Needed for the Fourth Tranche Payment for Most Firms



As a second check, there are 203 existing firms and 258 new firms in the experimental sample that are in business, and that have both a survey report of employment as well as the interviewers observation of the number of employees. I test whether there is any differential reporting effect by treatment status on this sub-sample by estimating:

$$\begin{aligned} &Survey\ report_i - InterviewerObservation_i \\ &\qquad = a + bAssignTreat_i + \ c * Region * Gender_i + \varepsilon_i \end{aligned}$$

The coefficient b is 0.91 (p=0.143) for the new enterprises, and 0.60 (p=0.466) for the existing enterprises. We can therefore not reject the null of no added difference in reporting with treatment group status. Although the point estimates are positive, they account for only 15 percent of the estimated treatment effect for new enterprises and 20 percent of the estimated treatment effect for existing enterprises. Thus even if selective over-reporting of employment by the treated is occurring in the survey, it only accounts for a small share of the overall treatment effect estimated. Finally note that any incentives to over-report employment should be lower in the third round, which occurs 12 to 18 months after individuals have received all funding from the program and yet we still see our treatment effects persist with this data.

Appendix 11.2 Total Jobs Created

Appendix Table A11.1 provides estimates of the total number of jobs created by the YouWin! program in its first round. To arrive at these estimates I take the experimental treatment effects and multiply by the number of firms in each group. For example, from Table 3 the impact on total employment for the new firms in the experimental treatment group is 5.2 workers per firm in round 3, which multiplied by 451 new firms in the treatment group, gives a total of 2359 additional jobs in round 3. Adding across new and existing firms then gives a total of 3,579 jobs created by the experimental sample at the time of the round 3 survey. I then bootstrap this process to account for the confidence intervals around the experimental point estimates, to arrive at a 95 percent confidence interval of (3061, 4161) for the jobs created by the experimental sample. I also use the propensity score estimates of the job creation for the national and zonal winners to likewise estimate the jobs created by this group, and aggregating the two gives a total of 7,027 jobs created by the winners in this competition.

Table A11.1 Total Employment in Winning Firms and Amount Attributable to Program

	Number of Firms	Total Employment in Winning Firms			Treatment Effect on Total Employment		
		Round 1	Round 2	Round 3	Round 1	Round 2	Round 3
Randomly selected winners	729	4588	7183	6858	1051	3411	3579
New Firms	451	2289	4209	4099	645	2711	2359
Existing Firms	278	2299	2974	2759	406	701	1220
National and Zonal winners	475	4439	6762	5870	1444	3366	3448
New Firms	118	744	1712	1273	320	1317	827
Existing Firms	357	3695	5050	4597	1125	2049	2620
All winners	1204	9027	13945	12728	2495	6777	7027

Notes:
Total Employment in Winning Firms is Survey Estimate of Average Employment Per Firm Multipled by the number of firms in that category.
Treatment Effects are experimental estimates for randomly selected winners, and propensity score matching estimates for national and zonal winners, multiplying the average impact per firm by the number of firms in that category.

This comparison of the treatment and control groups reflects the causal impact of the program on the difference in employment between the two groups. In order for this to reflect the overall impact on the economy, we need to make the following further assumptions: i) any wage job a YouWiN! winner leaves or doesn't take up is filled by someone else. If not, we should use the impacts on employment rather than on business ownership for the applicants themselves, which would lower the estimated employment creation by around 100 jobs; ii) YouWiN! firms do not destroy or generate jobs in other firms outside the experimental sample. If the winners compete with other

Nigerian firms and cause these firms to shut down or not expand as rapidly, the overall impact on employment is less. Conversely, if the firms provide complementary services that allow other firms to grow faster, the overall employment impact would be greater. I assume these two channels offset each other so that the first-order effect is zero here (some evidence to support the lack of large spillovers is discussed in the text and in appendix 20) and iii) the YouWiN! competition does not generate additional jobs through exciting non-winners to start businesses. It is possible the publicity and attention given to entrepreneurship motivates others to start businesses. Finally, note that this employment impact is the direct impact, and does not include any multiplier effects induced by the firms increasing demand for products of other firms, and by the firm owners and their employees increasing consumption of products made by other firms.

Appendix 11.3: Additional Employment Impacts

The table below shows the impacts on employment conditional on firm survival. Since these are conditioning on an outcome which is itself is affected by treatment, they should be considered as descriptive only. They show that even though the YouWiN! program has generated many more firms, the average number of employees per firm is still larger.

Table A11.2: Treatment-Control Difference in Employment Conditional on Survival

	New Firms			Existing Firms		
	Round 1	Round 2	Round 3	Round 1	Round 2	Round 3
Assigned to Treatment	0.189	4.459***	2.211***	0.907	3.229***	2.481***
	(1.109)	(0.472)	(0.570)	(0.824)	(0.791)	(0.757)
Sample Size	608	712	550	386	412	333
Control Mean	6.702	5.519	4.931	7.896	7.681	9.325

Notes: Robust standard errors in parentheses, *, **, *** indicate significance at the 10, 5, and 1 percent levels. Experimental estimates are OLS regression estimates and control for randomization strata.

Appendix table A11.3 shows impacts on other pre-specified employment measures. Treatment results in firms hiring more wage and salary workers and more casual and daily workers, but little change in unpaid workers. There is both more hiring and more firing of workers. The impact on the index of standardized z-scores is positive and significant.

Table A11.3: Impacts on Other Pre-Specified Employment Outcomes

	Wage & Salary Workers	Casual & Daily Workers	Unpaid Workers	Workers hired in last year	Workers fired in last year	Aggregate Employment Index
Panel A: New Firms						
First Follow-up	0.791***	0.423	-0.047	1.687***	0.190	0.237***
	(0.180)	(0.704)	(0.070)	(0.263)	(0.168)	(0.039)
Second Follow-up	3.627***	2.068***	0.115*	3.937***	1.339***	0.549***
	(0.231)	(0.303)	(0.059)	(0.429)	(0.292)	(0.035)
Third Follow-up	3.246***	1.555***	0.070	n.a.	n.a.	0.439***
	(0.296)	(0.319)	(0.072)			(0.034)
Control Mean: First follow-up	1.536	1.547	0.206	1.006	0.246	-0.075
Control Mean: Second follow-up	1.793	0.943	0.170	0.894	0.149	-0.180
Control Mean: Third follow-up	2.368	0.965	0.150			-0.149
Obs: First follow-up	992	982	980	995	995	1021
Obs: Second follow-up	1153	1149	1150	1153	1087	1181
Obs: Third follow-up	1007	1068	1067			1085
Panel B: Existing Firms						
First Follow-up	0.489	0.998*	0.222*	1.021**	0.064	0.201***
	(0.494)	(0.595)	(0.128)	(0.433)	(0.200)	(0.054)
Second Follow-up	2.656***	-0.111	0.030	2.258***	0.640**	0.267***
	(0.582)	(1.188)	(0.110)	(0.482)	(0.283)	(0.049)
Third Follow-up	2.961***	1.225***	-0.067	n.a.	n.a.	0.308***
	(0.509)	(0.417)	(0.103)			(0.053)
Control Mean: First follow-up	4.026	1.968	0.219	2.321	0.532	-0.105
Control Mean: Second follow-up	3.802	3.370	0.265	2.126	0.552	-0.136
Control Mean: Third follow-up	3.716	1.200	0.169			-0.160
Obs: First follow-up	422	418	418	423	423	432
Obs: Second follow-up	499	498	496	501	460	505
Obs: Third follow-up	450	472	470			477

Notes:
n.a. denotes question not asked in this survey round.
Robust standard errors in parentheses, *, **, *** indicate significance at the 10, 5, and 1 percent levels.
Regressions control for randomization strata.
Aggregate Employment Index is average of standardized z-scores of the owners' employment, the firm operating status, number of wage and salary workers, number of casual and daily workers, number of unpaid workers, and number of workers hired in the past year.

Appendix 12: Cost Effectiveness and Comparison to Cost per Job Generated in Other Studies

The text notes that the program spent $60 million to generate 7,027 jobs in the treated firms by the third round survey, for a cost of $8,538 per job created. The average wages of the jobs created in these firms is US$143 per month according to our survey reports, so the cost per job is equivalent to approximately 60 months of employment.

For this to represent the generation of additional jobs in the economy as a whole, and not just in the treated firms, requires several additional assumptions:: i) any wage job a YouWiN! winner leaves or doesn't take up is filled by someone else. If not, we should use the impacts on employment rather than on business ownership for the applicants themselves, which would lower the estimated employment creation by around 100 jobs; ii) YouWiN! firms do not destroy or generate jobs in other firms. If the winners compete with other Nigerian firms and cause these firms to shut down or not expand as rapidly, the overall impact on employment is less. Conversely, if the firms provide complementary services that allow other firms to grow faster, the overall employment impact would be greater. As discussed in the text, I do not find strong evidence for spillovers in either direction, and so I assume the first-order effect is zero here (while acknowledging they are difficult to measure); iii) any jobs the employees of winning firms would have otherwise taken up are now done by other people (the text notes that the majority of workers were either unemployed or students prior to working for the winning firms); and iv) the YouWiN! competition doesn't generate additional jobs through exciting non-winners to start businesses. It is possible the publicity and attention given to entrepreneurship motivates others to start businesses. Finally, note that this employment impact is the direct impact, and does not include any multiplier effects induced by the firms increasing demand for products of other firms, and by the firm owners and their employees increasing consumption of products made by other firms. It also assumes that the government funding used for this program did not lower employment in the process of being collected, as might be the case with distortionary taxes. Since the funding from the program is largely from oil revenues, this does not seem the case here.

As the above set of assumptions illustrates, cost-effectiveness calculations require many additional assumptions to those needed for the experimental estimates themselves, and so one should exercise caution in interpreting these cost per job numbers.

Nevertheless, the fact that I find impacts on employment at all means that the cost per job created compares favorably with many job creation policy efforts in developing countries, which have struggled to find significant effects on employment. Most studies of finance and training programs for microenterprises have at best found impacts on generating self-employment, but often offset by a reduction in wage employment of the potential entrepreneur (McKenzie and Woodruff, 2014), with little discernable effect on paid employment (Grimm and Paffhausen, 2015). Management consulting services to larger firms have also often not led to jobs, with the one study that has found impacts only finding these over the medium-term with administrative data on formal employment (Bruhn et al, forthcoming). Blattman and Ralston (2015) also highlight the limited success of microfinance, insurance, and several other policies in creating employment.

Appendix Table 12 gives examples of the cost per job created from other impact evaluations in developing countries, for studies in which cost and employment impact data were available. Many studies do not provide cost data, limiting the set of studies for which this comparison can be done. For the vocational training and wage subsidy studies, the studies look at whether the program increases the employment likelihood of the person receiving the training or subsidy. For the management consulting, small grants to microenterprises, and business training interventions, the studies also include whether the firm hires paid employees. Note that the absence of an impact on paid employment does not necessarily mean these interventions are ineffective: the may also have impacts on the earnings of workers and firms, on a shift to formal employment, and on other outcomes of policy interest.

Table A12: Examples of Cost per Job Generated in other Randomized Controlled Trials in Developing Countries

Intervention Type	Study	Country	Employment Impact	Cost Per Treated Unit (USD)	Cost Per Job Created (USD)
Vocational Training	Hirshleifer et al. (2015)	Turkey	0.02 (n.s.)	1619	80950 (n.s.)
	Attanasio et al. (2011)	Colombia	0.068 (females), 0.013 (males, n.s.)	750	11029 (females), 57692 (males, n.s.)
Wage Subsidies	Groh et al. (2014)	Jordan	0.015 (n.s.)	571	38100 (n.s.)
Management consulting	Bloom et al. (2013)	India	-1.28 (n.s.)	75,000	no creation
	Karlan et al. (2015)	Ghana	0.047 (n.s.)	1125	23936 (n.s.)
	Bruhn et al. (2013)	Mexico	0.52 (one year, n.s.); 4.43 (admin long-run)	11856	2676 (long-run), 22800 (one year, n.s.)
Small grants to microenterprises	De Mel et al. (2012)	Sri Lanka	-0.03 (n.s.)	100-200	no creation
	Karlan et al. (2014)	Ghana	-0.169 (n.s.)	133	no creation
Business training	Karlan and Valdivia (2011)	Peru	0.017 (n.s.)	n.a.	n.a.
	Valdivia (2015)	Peru	-0.06 (n.s.)	337	no creation
	Drexler et al. (2014)	Dominican Republic	-0.02 (rule of thumb, n.s.), 0.05 (n.s. standard training)	21	no creation, 420 (standard training, n.s.)

Notes:
n.s. denotes not statistically significant, n.a. denotes cost data not available. Most studies costs are direct costs only, and do not include program implementation costs.
Bruhn et al. (2013)'s long-run estimates are only for the subsample of firms they could match to administrative data, and captures formal employment only.

The above table considers only evidence from randomized controlled trials for which concerns about selection are lowest. In a recent working paper, Fafchamps and Quinn (2016) report that US$40,000 in prizes created 80 new wage jobs, resulting in a very low cost per job created of $500 (they do not provide details of the administrative costs of running their competition, so these costs are excluded). These results suggest the potential for business plan competitions with smaller prize amounts to also be able to generate jobs in a very cost effective manner. Two caveats are that this impact comes from comparing winners to runners-up, resulting in a potential upward bias in jobs impact if the winners have better growth prospects; and that the impact is only measured in the first six months after the grants were given. De Mel et al. (2014) find that grants combined with training had very large impacts on start-up enterprises in Sri Lanka over 4 and 8 month horizons, but that these impacts disappear when measured over longer horizons of 16 and 25 months. It will therefore be important to see whether these smaller-scale competitions have effects that last.

Appendix 13: Impacts on Different Innovative Activities

Innovation is an index of 12 different types of innovation, including product, process, marketing, pricing, quality control, and use of the internet. Table 3 showed a 14 to 22 percentage point increase in innovative activities for experimental winners. However, some of this reflects that firms need to be in business in order to innovate, and that the program had large impacts on start-up and survival. In appendix Table A13, I look descriptively at the treatment-control difference conditional on the business operating. The new applicants are innovating more in multiple dimensions, introducing new products, processes, pricing methods, quality control systems, using the internet, and using new channels for selling goods. Existing firm applicants are more likely to have introduced a new product, introduced a new channel for selling, and to be using the internet. These results suggest that winning firms are increasing the variety of products they produce through innovation.

The survey asks firms about the new products they have introduced through innovation. Among the winning firms in the treatment group, 68 percent said the new product was new for the firm, but other firms in the same city sell something similar; 20 percent said it was new for the city but available elsewhere in Nigeria; 10 percent said it was new for Nigeria but available elsewhere in the world; and 1.5 percent said it was new for the world. So the most common type of innovation is the most typical type in developing countries, of copying and introducing locally products available elsewhere.

Table A13: Treatment-Control Difference in Innovative Activities Conditional on Operating

	New Firms			Existing Firms		
Outcome	Control Mean	Round 2	Round 3	Control Mean	Round 2	Round 3
Introduced a new product	0.377	0.183***	0.016	0.348	0.096**	0.056
		(0.037)	(0.041)		(0.049)	(0.055)
Improved existing product or service	0.582	0.078**	0.038	0.428	0.071	0.011
		(0.038)	(0.065)		(0.048)	(0.070)
Introduced new or improved process	0.508	0.131***	0.071*	0.406	0.079	0.023
		(0.037)	(0.038)		(0.049)	(0.054)
Introduced new design or packaging	0.563	0.164***	0.109***	0.497	0.069	0.063
		(0.038)	(0.041)		(0.048)	(0.055)
Introduced new channel for selling goods	0.525	0.145***	0.114***	0.473	0.086*	0.126**
		(0.038)	(0.041)		(0.049)	(0.054)
Introduced new method for pricing	0.612	0.147***	0.116***	0.535	0.061	0.009
		(0.037)	(0.040)		(0.048)	(0.053)
Introduced new method of advertising	0.656	0.173***	0.107***	0.543	0.033	-0.014
		(0.036)	(0.041)		(0.047)	(0.052)
Changed way work organized in firm	0.585	0.199***	0.065	0.428	0.071	0.042
		(0.037)	(0.040)		(0.048)	(0.055)
Introduced new quality control standards	0.481	0.106***	0.100**	0.358	0.002	0.032
		(0.037)	(0.041)		(0.050)	(0.053)
Licensed a new technology	0.186	0.070**	-0.003	0.126	0.043	-0.088*
		(0.029)	(0.026)		(0.040)	(0.045)
Obtained new quality certification	0.126	0.095***	-0.007	0.053	-0.020	0.023
		(0.024)	(0.021)		(0.032)	(0.028)
Uses internet	0.699	0.118***	0.105***	0.650	0.105**	0.050
		(0.034)	(0.037)		(0.042)	(0.045)

Notes: robust standard errors in parentheses, *, **, *** indicate significance at the 10, 5, and 1 percent levels respectively.
Each row contains the OLS coefficient on being assigned to treatment for a regression where the outcomes is a measure of innovation, conditional on the business being in operation. Randomization strata are controlled for.

Appendix 14: Robustness of Sales and Profit Impacts

Appendix Figures A14.1 and A14.2 shows the cumulative distribution functions of round 3 sales and profits by treatment status. There are two important points to note from these CDFs. The first is that the treatment distribution lies to the right of control distribution below the 90th percentile in all four graphs. This is reflected in the positive quantile treatment effects for profits reported in the text and in Appendix Figure A14.3 for sales. However, the second point to note is that the upper tails of the distribution are quite long, and the distributions cross each other several times at the upper tail. This long tail makes impacts estimated in levels highly sensitive to the few observations at the top of the distribution, and greatly reduces statistical power for estimating the impact on the level of profits and sales (for example, the standard deviation of profits for the control group among new firms is 1.2 million, relative to a mean of 174,000).

There are several approaches one can take to increase power or to examine outcomes that are less sensitive to the few firms at the top. The pre-analysis plan specified truncating the distributions at the 99th percentile, taking the inverse hyperbolic sine of profits (which is similar to a log transformation but can allow for zeros and negative values), and using an aggregate index of standardized z-scores of different measures. Truncating at the 99th percentile only partly deals with the issue, and still results in a relatively large coefficient of variation, rendering power low.

Appendix Table A14.1 then considers robustness to several alternative transformations (not pre-specified). There are three different transforms that are commonly used to approximate a logarithmic transformation in cases where the outcome can take zero values. The first (pre-specified) is the inverse hyperbolic sine. Second, I consider the transform log(x+1), where x is profits or sales respectively. Third, I consider the quintic root of profits or sales. I then also consider two binary transforms of the data: having profits of at least 100,000 Naira per month, or at least 200,000 Naira per month; and sales of at least 200,000 Naira per month, and at least 500,000 Naira per month. The treatment effects on sales and profits are large and statistically significant under all these alternative transformations. Combined with the quantile treatment effects analysis shown in the paper, and with the impact on the standardized z-scores, I view this evidence as showing that treatment did have positive and significant impacts on profits and sales at the time of the last follow-up survey, and that the lack of significance of the impact on the level of truncated profits is due to sensitivity to the uppermost tail.

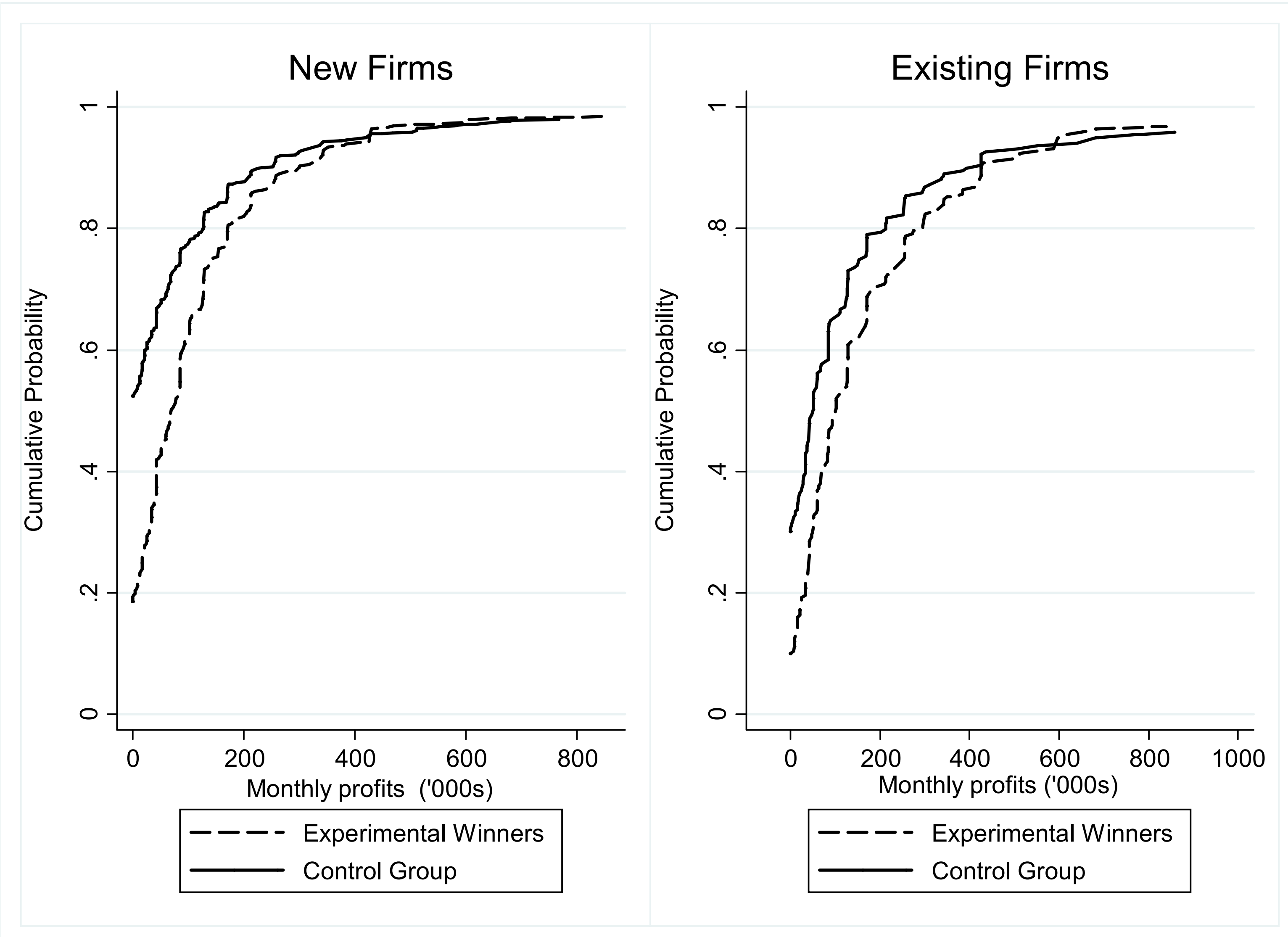
A further way to improve power is to pool together the round 2 and round 3 waves, thereby calculating an average impact over the two rounds. McKenzie (2012) shows this can be particularly helpful in improving power for firm outcomes like profits and sales which are not strongly autocorrelated. Appendix Table A14.2 provides the pooled estimates, which are all statistically significant.

Appendix Table A14.3 examines the impact of winning on several other measures of profits and sales. These include the impacts on several measures included in the aggregate index of sales and profits in Table 4 (untruncated monthly sales and untruncated monthly profits, annual sales, whether the firm says sales are higher than a year ago, and profits in the best month), along with several measures that were only asked in the first two survey rounds since they proved difficult for firms to answer (number of customers served in the past week, sales of the main product, and mark-up profits). Many of these measures are subject to the same upper tail issues as discussed above. Positive treatment effects are seen on having more customers (for round 2, new firms); annual sales; having sales higher than a year ago; and profits in the best month of the last 12.

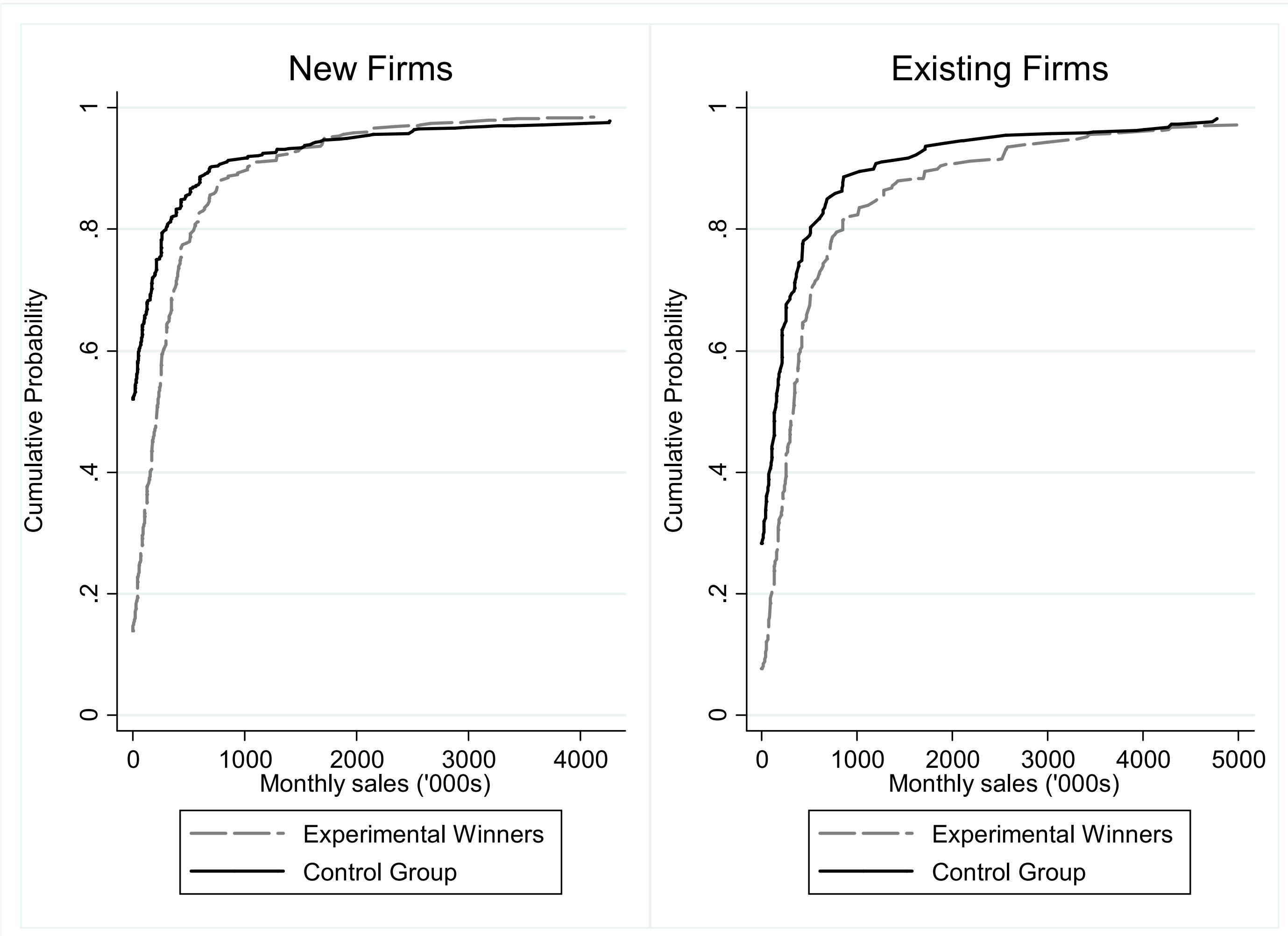
A concern with any program involving some business training or improvements in record-keeping is that it may lead to changes in the accuracy of information being reported, even if the underlying business financial position does not change. If businesses systematically under- or over-state sales and profits, this will lead to a bias in the measured treatment effect. If this is the case, we might expect changes in the number of inconsistencies or errors in reporting of profits and sales. I consider four reporting errors: a) total sales in the last month exceed total sales in the year to date so far; b) profits in the last month exceed sales in the last month; c) profits in the best month of the year are less than profits in the last month; and d) revenues in the last month from the main product[5] (calculated as price per unit times number of units sold) exceed reported total revenues for the last month. The control group made about 0.8 errors on average in the first round, and 0.12 on average in the second round, with a large part of this drop reflecting better interviewer training on how to ask about the main product. Appendix Table A14.4 shows that there is no differential treatment effect for existing firms in the number of errors made in either round. Among new enterprises who are in business, treated firms make 0.04 fewer errors than control firms, which is marginally significant at the 10 percent level in the second round. This difference is small in magnitude, and the sharpened q-value for this result taking account of testing four outcomes in this table is 0.479. Therefore it does not appear that treatment is resulting in large differences in reporting behavior.

[5] This variable was not asked in round 3, so we focus only on the first two rounds.

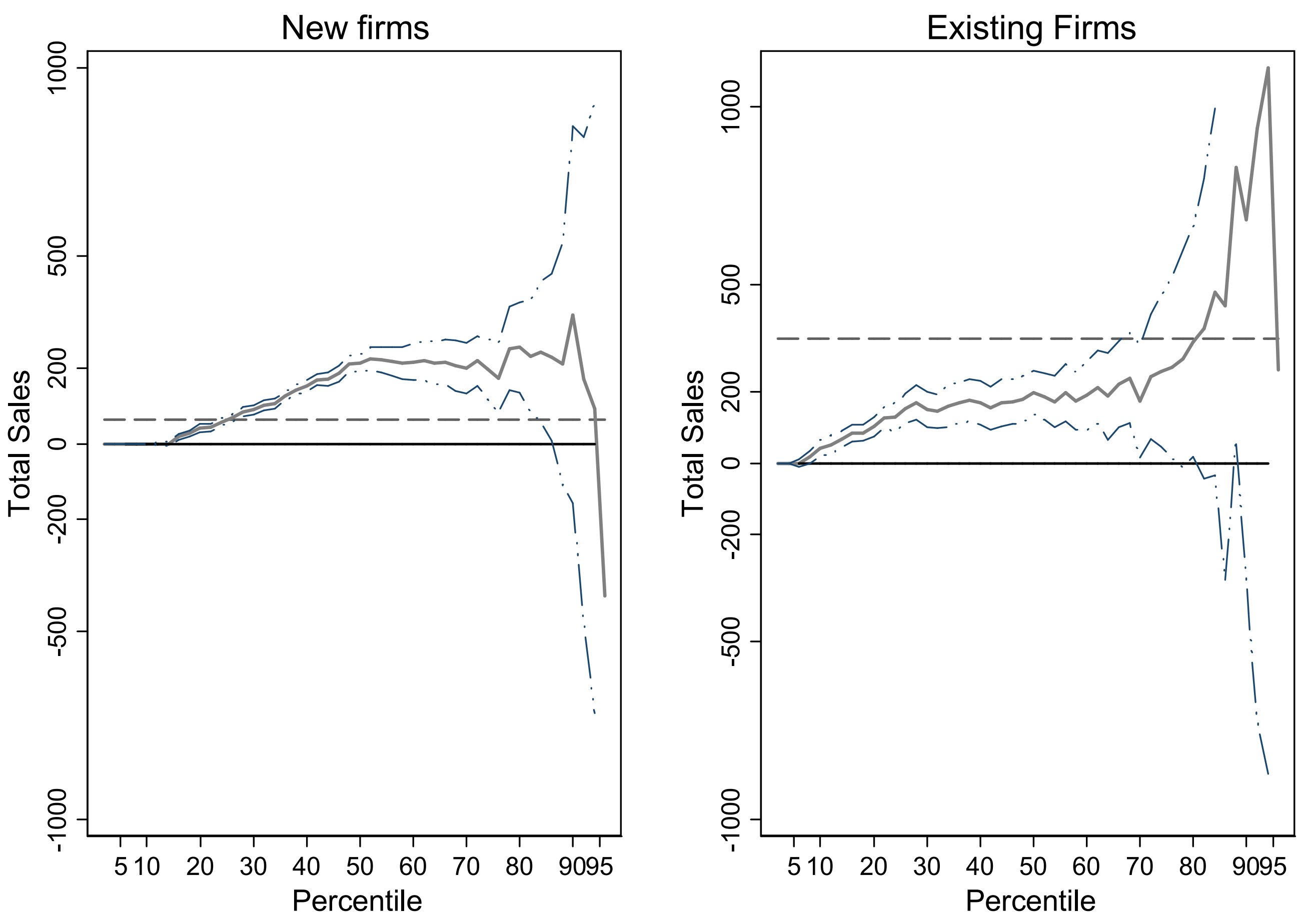
Appendix Figure A14.1: Cumulative Distribution Functions for Profits in Round 3 by Treatment Status



Appendix Figure A14.2: Cumulative Distribution Functions for Sales in Round 3 by Treatment Status



Appendix Figure A14.3: Quantile Treatment Effects for Sales in Round 3



Notes: Sales are in thousands of real Naira per month. Notes: New firms and Existing firms refer to status at time of application. Round 3 is three years after application and 12-18 months after all grants have been received. 95 percent confidence interval shown around quantile treatment effect. Dashed line indicates OLS point estimate.

Appendix Table A14.1: Robustness of Third Round Profit and Sale Impacts to Different Transformations

	Profits					Sales				
	Inverse Hyperbolic	Log (x+1)	Quintic Root	Above 100K	Above 200K	Inverse Hyperbolic	Log (x+1)	Quintic Root	Above 200K	Above 500K
Panel A: New Firms										
Assignment to Treatment	3.962***	1.461***	0.801***	0.159***	0.065***	2.332***	2.076***	1.142***	0.258***	0.081***
	(0.346)	(0.142)	(0.080)	(0.030)	(0.024)	(0.182)	(0.165)	(0.098)	(0.031)	(0.026)
Sample Size	1063	1063	1063	1063	1063	1063	1063	1063	1063	1063
Control Mean	5.775	2.171	1.226	0.226	0.125	2.993	2.666	1.531	0.272	0.144
Panel B: Existing Firms										
Assignment to Treatment	2.580***	1.060***	0.578***	0.164***	0.088**	1.618***	1.474***	0.854***	0.250***	0.117***
	(0.464)	(0.202)	(0.114)	(0.046)	(0.040)	(0.249)	(0.227)	(0.137)	(0.045)	(0.041)
Sample Size	469	469	469	469	469	468	468	468	468	468
Control Mean	8.565	3.273	1.841	0.347	0.210	4.524	4.033	2.287	0.429	0.215

Notes: Robust standard errors in parentheses, *, **, and *** denote significance at the 10, 5, and 1 percent levels respectively. Regressions control for randomization strata.

Appendix Table A14.2: Pooled Round 2 and 3 estimates

	New Firms				Existing Firms			
	Truncated Sales	Truncated Profits	Inverse Hyperbolic Sine Profits	Aggregate Index of Sales and Profits	Truncated Sales	Truncated Profits	Inverse Hyperbolic Sine Profits	Aggregate Index of Sales and Profits
Pooled Second and Third Round Effect	185.142***	45.608***	4.062***	0.235***	336.931***	50.434*	2.357***	0.223***
	(60.642)	(14.550)	(0.264)	(0.032)	(110.685)	(30.255)	(0.338)	(0.052)

Notes: Robust standard errors in parentheses, clustered at the firm level. *, **, *** indicate significance at the 10, 5, and 1 percent levels.
Experimental estimates are ITT estimates and control for randomization strata.
Sales and Profits are in 1000s of real Naira per month.
Aggregate index of outcomes includes monthly sales, truncated monthly sales, annual sales, sales higher than one year ago, monthly profits, truncated monthly profits, profits in the best month, and inverse hyperbolic sine of profits.

Appendix Table A14.3: Impacts on Other Pre-specified Sales and Profits Outcome Measures

	Number of Customers in week	Untruncated Monthly Sales	Annual Sales	Sales are higher than year ago	Untruncated Monthly Profits	Profits in best month	Sales of main product	Mark-up profit on main product	Aggregate outcome index
Panel A: New Firms									
First Follow-up	9.748	119.381	-248.263	-0.053*	-49.163	-19.114	4555.891*	1292.952*	0.044
	(6.098)	(97.599)	(162.620)	(0.031)	(36.547)	(28.329)	(2763.075)	(714.581)	(0.042)
Second Follow-up	14.109**	148.377	1765.721***	0.205***	39.346	111.898***	121.812***	42.756***	0.266***
	(6.362)	(134.477)	(483.571)	(0.031)	(30.743)	(21.165)	(34.218)	(13.125)	(0.035)
Third Follow-up		-26.983	802.723**	0.203***	-50.874	83.643***			
		(129.783)	(381.820)	(0.031)	(59.576)	(31.141)			
Control Mean: First follow-up	27.965	277.280	1271.742	0.369	195.740	188.660	2680.548	775.056	-0.015
Control Mean: Second follow-up	32.600	502.419	2022.581	0.393	139.112	124.074	170.514	61.608	-0.087
Control Mean: Third follow-up		528.777	2197.340	0.341	174.143	154.308			
Obs: First follow-up	989	995	995	995	995	995	989	954	995
Obs: Second follow-up	1152	1151	1069	1151	1150	1071	1142	1141	1156
Obs: Third follow-up		1063	925	1063	1063	927			
Panel B: Existing Firms									
First Follow-up	5.992	40.778	286.407	0.082*	-10.975	24.431	-733.669	671.735	0.066
	(10.160)	(89.075)	(386.195)	(0.043)	(57.133)	(55.678)	(4823.938)	(1127.670)	(0.059)
Second Follow-up	20.329	302.230	1874.552*	0.172***	128.944*	131.023**	100.325	33.633	0.205***
	(12.766)	(401.365)	(1009.407)	(0.038)	(76.449)	(57.770)	(62.374)	(23.394)	(0.058)
Third Follow-up		450.729**	2068.898**	0.095**	70.043	74.256			
		(196.860)	(847.843)	(0.046)	(61.668)	(59.610)			
Control Mean: First follow-up	45.473	519.907	2697.286	0.684	271.504	337.512	9111.524	1773.691	-0.035
Control Mean: Second follow-up	42.167	982.920	4770.229	0.664	225.071	327.765	313.702	122.960	-0.103
Control Mean: Third follow-up		509.975	3367.593	0.516	196.047	296.349			
Obs: First follow-up	420	423	423	423	423	423	420	411	423
Obs: Second follow-up	500	497	458	497	497	458	496	496	501
Obs: Third follow-up		468	409	470	469	409			

Notes: Robust standard errors in parentheses, *, **, *** indicate significance at the 10, 5, and 1 percent levels.
Experimental estimates are ITT estimates and control for randomization strata.
Aggregate index of outcomes includes monthly sales, truncated monthly sales, annual sales, sales higher than one year ago, monthly profits, truncated monthly profits, profits in the best month, inverse hyperbolic sine of profits, number of customers, sales of main product, and mark-up profit.

Appendix Table A14.4: Impact on Profit and Sales Reporting Errors

	Existing Firms in Operation		New Firms in Operation	
	First Round	Second Round	First Round	Second Round
Experimental Treatment Effect	0.018	0.004	-0.044	-0.048*
	(0.081)	(0.038)	(0.068)	(0.027)
Sample Size	384	413	610	706
Control Mean	0.780	0.115	0.853	0.142

Notes: robust standard errors in parentheses, *, **, *** indicate significance at the 10, 5, and 1 percent levels respectively

Dependent variable is the total number of reporting errors out of 4 made (monthly sales>annual sales, profits>sales, profits in best month<profits in last month, and sales of main product>total sales). Note main product sales not asked in round 3.

Appendix 15: Different Mechanisms Leading to Impact

Appendix Table A15 examines the impacts of winning on several mechanisms that potentially affect firm productivity (A) and entrepreneurial skills (E). The first column examines entrepreneurial self-efficacy, a measure of the owner's self-confidence in their ability to carry out 12 business-related actions such as "estimate customer demand for a new product", and "identify good employees who can help the business grow". The mean owner is very confident in their ability to do 5 out of 12 tasks, and winning the competition has no impact on this measure. This measure captures a combination of actual skill and confidence, and suggests little change in E.

The second and third columns look at mentoring and use of business networks. Mentoring measures whether the business owner has a mentor they talk to about business matters. This is significant for existing firms in the unconditional regressions, but with one exception, not in the results conditional on operating a business. Network measures the number of other firm owners the business owner discusses business matters with, which is again not significant in the conditional regressions.

Column 4 examines the impact on business practices, as measured by the proportion of 22 business practices employed using the measures set out in McKenzie and Woodruff (2015). There is a positive and significant impact on the unconditional estimates, which code practices as zero for those not operating a business. However, conditional on operating, there are no significant impacts for existing firms, while for new firms the impacts are significant but small in magnitude (equivalent to 1.3 more practices out of 22 being done). The business training literature surveyed by McKenzie and Woodruff (2015) suggests that this magnitude of change in business practices would be predicted to result in only a 0.2 percent increase in the likelihood of operating a business, and a 4 percent increase in business profits. This suggests that any impact on business practices is not likely to explain much of the treatment effects found in this paper.

Column 5 examines formality, measured in terms of whether the firm has a registered business name, municipal license, and income tax registration. We see a large increase in this measure, which is consistent with the winners needing to register in some form to receive grant payments. The existing global evidence suggests that formality per se does not have measureable impacts on firm productivity or performance for most firms (Bruhn and McKenzie, 2014).

Appendix Table A15: Impact on A and E

	Entrepreneurial Self-Efficacy	Has a Mentor	Number in Business Network	Business Practices Index	Firm is Formal
Panel A: New Firms					
Unconditional Experimental Impacts:					
First-Follow-up	0.226	0.206***	0.563***	0.152***	0.084***
	(0.179)	(0.032)	(0.216)	(0.025)	(0.021)
Second Follow-up	0.045	0.381***	1.709***	0.339***	0.382***
	(0.175)	(0.028)	(0.210)	(0.021)	(0.028)
Third Follow-up	0.300*	0.334***	1.711***	0.358***	0.368***
	(0.180)	(0.032)	(0.235)	(0.022)	(0.032)
Impacts Conditional on Business in Operation					
First-Follow-up	0.071	0.044	-0.371	-0.013	0.075**
	(0.218)	(0.033)	(0.293)	(0.017)	(0.030)
Second Follow-up	-0.033	0.063**	0.411*	0.050***	0.323***
	(0.206)	(0.025)	(0.247)	(0.012)	(0.034)
Third Follow-up	0.204	0.041	0.438	0.060***	0.295***
	(0.208)	(0.033)	(0.275)	(0.014)	(0.037)
Control Mean: First follow-up	4.765	0.429	2.029	0.406	0.067
Control Mean: Second follow-up	5.523	0.445	1.879	0.409	0.114
Control Mean: Third follow-up	5.513	0.443	1.964	0.341	0.126
Obs: First follow-up	973	995	992	995	995
Obs: Second follow-up	997	1071	1071	1071	1071
Obs: Third follow-up	859	857	857	927	857
Panel B: Existing Firms					
Unconditional Experimental Impacts:					
First-Follow-up	0.003	0.114***	0.762**	0.081***	0.082**
	(0.259)	(0.040)	(0.347)	(0.027)	(0.041)
Second Follow-up	0.015	0.150***	1.082***	0.133***	0.310***
	(0.257)	(0.038)	(0.337)	(0.026)	(0.043)
Third Follow-up	-0.245	0.085**	0.828**	0.183***	0.242***
	(0.238)	(0.043)	(0.340)	(0.032)	(0.050)
Impacts Conditional on Business in Operation					
First-Follow-up	0.006	0.046	0.483	0.015	0.065
	(0.269)	(0.036)	(0.361)	(0.018)	(0.045)
Second Follow-up	-0.112	0.027	0.544	0.017	0.277***
	(0.267)	(0.033)	(0.348)	(0.014)	(0.047)
Third Follow-up	-0.190	-0.013	0.380	0.003	0.204***
	(0.251)	(0.040)	(0.346)	(0.019)	(0.053)
Control Mean: First follow-up	5.135	0.726	3.095	0.687	0.184
Control Mean: Second follow-up	5.507	0.719	3.059	0.676	0.226
Control Mean: Third follow-up	6.034	0.736	3.356	0.559	0.307
Obs: First follow-up	418	423	423	423	423
Obs: Second follow-up	448	458	458	458	458
Obs: Third follow-up	392	372	372	409	372

Notes: Control means shown are unconditional means
Business Practices is the proportion of 22 business practices employed.
Entrepreneurial Self-efficacy is the number of 9 activities that the individual is very confident they can c
Robust standard errors in parentheses, *, **, *** indicate significance at the 10, 5, and 1 percent levels.
Regressions control for randomization strata.

Appendix 16: Further Evidence on Capital

Appendix 16.1: Did the Grants Crowd Out Informal Finance?

Table 5 shows little impact of winning on the receipt of formal loans or outside equity financing. Appendix Table A16.1 shows the impact on informal loan financing (from moneylenders, family, or friends). This outcome was not pre-specified, but suggested by a referee. This form of borrowing is more common than formal loans, but still taken by very few firms. There is no significant impact on informal borrowing for existing firms, nor for new firms in two out of three rounds. There is a significant positive impact on receipt of informal loans for new firms in round 3. However, the sharpened q-value that takes account of testing 6 different outcomes in this table is 0.124 for this outcome. The evidence here shows informal borrowing is relatively rare, and does not appear to have been crowded out by the receipt of the grant.

Appendix Table A16.1: Impact on Informal Loan Financing

	New firms			Existing firms		
	Round 1	Round 2	Round 3	Round 1	Round 2	Round 3
Assigned to Treatment	-0.015	0.011	0.045**	-0.013	-0.031	0.005
	(0.013)	(0.016)	(0.019)	(0.022)	(0.027)	(0.030)
Sample Size	995	1071	857	423	458	372
Control Mean	0.048	0.065	0.051	0.058	0.104	0.086

Notes:
Robust standard errors in parentheses. *, **, and *** indicate significance at the 10, 5, and 1 percent levels respectively. Regressions also control for randomization strata.
Informal loans include loans from moneylenders, family, and friends.

Appendix 16.2: Heterogeneity with Respect to Importance of Capital

I contend in the paper that the main effect of the business plan competition is to allow firms to overcome credit constraints, using the capital granted to purchase more capital, produce a wider variety of products, and hire more labor. A natural question is then whether we see heterogeneity in impacts by how binding these credit constraints are to begin with. I consider here two measures of heterogeneity, both suggested by a referee (and therefore not pre-specified).

The first measure is based on the capital intensity of the industry. We might expect the impact of the grant to be greater in industries that are more capital intensive. However, while this may be true for a random sample of firms from each industry, it is less clear that we should expect the selected sample of semi-finalists in the program to necessarily exhibit the same pattern. A second complication is how to measure the capital intensity of the industry. Nigeria does not have any recent census data on firms available to characterize the capital intensities of industries within the country, and if credit constraints are pervasive, then it is unclear that the realized capital intensities necessarily reflect the relative capital intensities that would prevail in the absence of these constraints. I therefore use the classification of Guerrieri and Acemoglu (2008) for U.S. firms, based on the idea that these firms are less likely to be constrained. Based on their binary classification of low capital intensity industries, I classify firms in personal services (e.g. hair and beauty), repair services, professional services, and construction services (e.g. plumbing and electrical work) as low capital intensity. I do this on the basis of the self-classification of the firm into industry at the time of baseline application, which provides this information for all applicants. 28 percent of new applicants and 37 percent of existing applicants are in low capital-intensity industries. I then estimate the following equation separately for new and existing firms, for key round 3 outcomes:

$$Outcome_i = a + b * AssignTreat_i + c * Region * Gender_i + d * LowIntensity_i + e * AssignTreat_i * LowIntensity_i + \varepsilon_i \quad \text{(C1)}$$

The prediction is then that e<0, that is, the impact of the grants should be lower in low capital intensity industries. Panel A of Table A16.2 examines this prediction. The prediction is born out in terms of sign, with the interaction negative for all five outcomes considered for new firms, and for four out of five outcomes for existing firms. However, only the impact on new firms for having 10 or more workers is significant at the 10 percent level, and the sharpened q-value is 1 still for this outcome.

The second measure is a self-assessed measure which comes from the baseline datasheet collected at the time of submitting their full business plan. In this sheet, I asked them what the major challenges to running your own business are, with "lack of funds to start a business" being one of the responses. 95% of new firm applicants and 67% of existing firm applicants say this is a major constraint. I interact a dummy variable for saying finance is a constraint with treatment, and include this together with the interaction in an analogous regression to equation (C1). The prediction is then that this interaction should be positive, that is, those for whom financing is a constraint should benefit more from the grant. Panel B shows the result. For new firms we see all five coefficients are positive, and the coefficients for impacts on total employment (p=0.002) and having 10 or more workers (p=0.001) are strongly significant and survive corrections for multiple testing (sharpened q-values are 0.009 when considering the 10 outcomes). In contrast, none of the interactions are significant for existing firms, and four out of five are in fact negative. Ex post one might speculate that for an existing firm, saying that finance is not a constraint is an indication of less ambition to grow.

Taken together, I view these results as offering modest support for the idea that more capital-constrained firms benefit more from the grants, at least amongst the new applicants. But the strongest heterogeneity comes from a case where almost everyone says they are constrained. The quantile treatment effects are certainly consistent with their being gains across the distribution and the impacts not being concentrated on one small subset of firms. It may well be that amongst the selected sample of those who apply for the competition, and get judged as high-scoring contestants, most firms face constraints and there is therefore insufficient heterogeneity in whether or not capital is a constraint to examine this more deeply. Alternatively, improvements in the measurement of capital constraints in future work may be able to better distinguish which firms are relatively more constrained than others.

Appendix Table A16.2: Heterogeneity of Impacts with respect to proxies for importance of capital constraints

	New Firms					Existing Firms				
	Operates Firm	Total Employment	10+ Workers	IHS Profits	Profits and Sales Index	Operates Firm	Total Employment	10+ Workers	IHS Profits	Profits and Sales Index
Panel A: Heterogeneity with respect to Capital Intensity of Industry										
Assigned to Treatment	0.384***	5.373***	0.254***	4.117***	0.186***	0.197***	4.331***	0.228***	2.609***	0.230**
	(0.028)	(0.566)	(0.033)	(0.404)	(0.050)	(0.042)	(0.848)	(0.051)	(0.601)	(0.094)
Assigned to Treatment*Low Capital Industry	-0.041	-0.635	-0.100*	-0.587	-0.075	-0.007	0.224	-0.052	-0.096	-0.051
	(0.054)	(0.924)	(0.059)	(0.768)	(0.088)	(0.064)	(1.347)	(0.084)	(0.964)	(0.149)
Sample Size	1085	1044	1044	1063	1063	477	461	461	469	470
Panel B: Heterogeneity with respect to saying "lack of funds to start a business is main challenge"										
Assigned to Treatment	0.289**	0.005	-0.128	3.863**	-0.047	0.210***	4.226***	0.291***	3.305***	0.167
	(0.125)	(1.667)	(0.110)	(1.551)	(0.160)	(0.052)	(1.113)	(0.076)	(0.796)	(0.127)
Assigned to Treatment*Lack of Funds	0.088	5.413***	0.369***	0.120	0.223	-0.020	-1.341	-0.061	0.169	-0.145
	(0.127)	(1.718)	(0.113)	(1.587)	(0.165)	(0.061)	(0.941)	(0.057)	(0.847)	(0.110)
Sample Size	1085	1044	1044	1063	1063	477	461	461	469	470

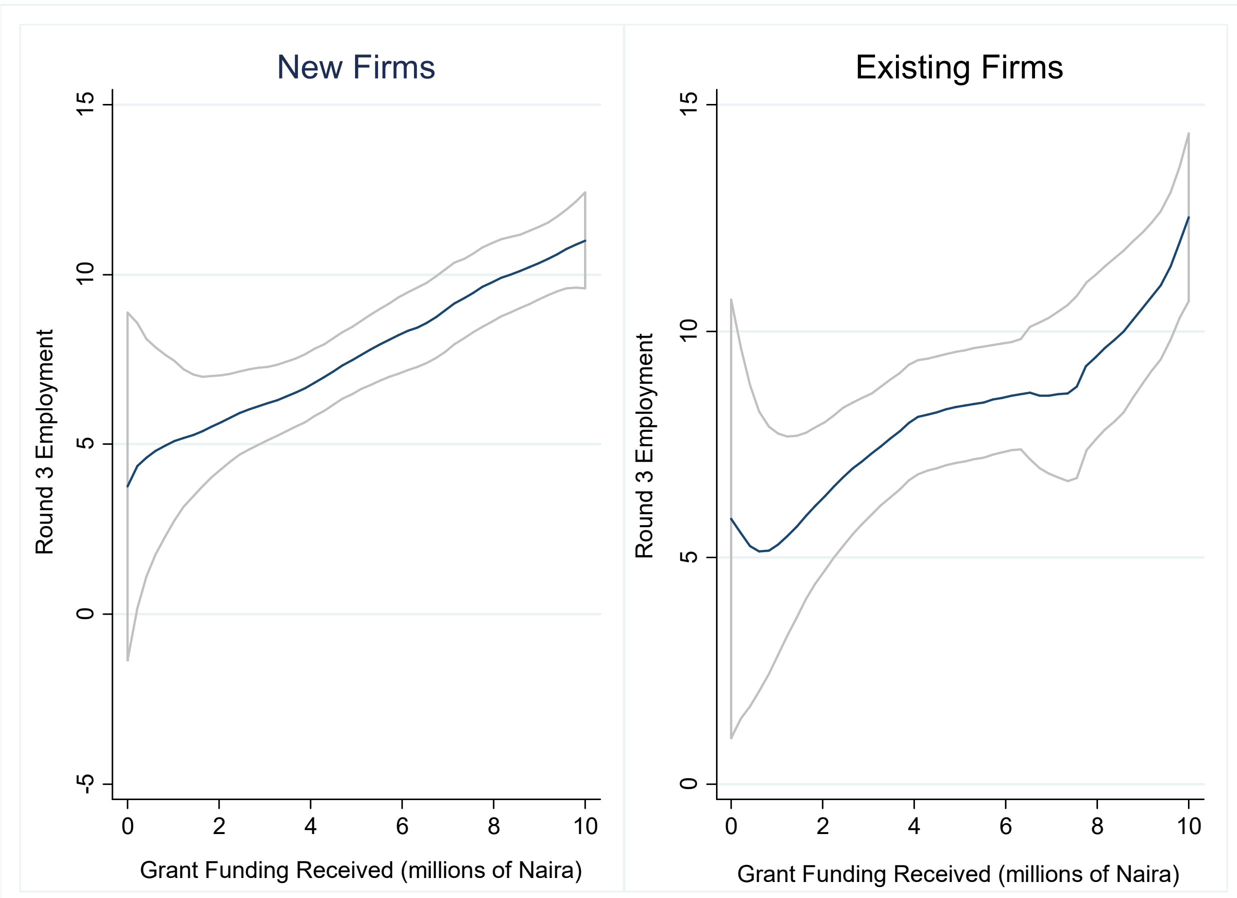
Notes:
Robust standard errors in parentheses. *, **, and *** indicate significance at the 10, 5, and 1 percent levels respectively.
Existing and new refer to firm status at time of application; low capital intensity refers to sectors which have low capital intensity in the U.S. as determined by Guerrieri and Acemoglu (2008); lack of funds to start a business is major challenge is response of individual on baseline data survey.
Regressions control for randomization strata and for level effect of the interacting term.

Appendix 16.3: Heterogeneity with respect to amount of funding received

As shown in Appendix 3, while the median experimental winner received 7.7 million Naira ($49,000) in funding, there was considerable variation in the amount received, with the 10th percentile receiving 3.4 million Naira ($22,000) and the 90th percentile receiving the maximum of 10 million Naira ($64,000). This raises the question of whether those winners who received more funding were able to grow larger.

Appendix Figure A16 answers this descriptive question by plotting local linear regressions between employment in round 3 and the amount of funding received. We see a positive, and approximately linear, relationship between the amount of funding received and employment among these winners.

Appendix Figure A16: Relationship between Funding Level and Round 3 Employment for Experimental Winners



Note: local linear regression curves shown using Epanecknikov kernel with Stata's default bandwidth choice. 95 percent confidence intervals also displayed.

The amount of grant received was not randomly decided amongst winners, but reflects their business plan request, and the assessment of the experts as to whether the budget in their plan needed modification. It also reflects whether or not applicants met the triggers to receive all four tranche payments. To explore this association further, in appendix Table A16.3, I consider only the firms which received all four tranche payments and regress round 3 outcomes for the experimental treatment group on the grant amount received. We see a strong and statistically significant positive association between the amount received and total employment, having 10 or more employees, and the profits and sales index for both new and existing firms. I then control for selection on observables by controlling for the set of baseline characteristics set out in Table 1 (which includes the application and business plan scores). This does not change the magnitudes or statistical significance. Under the assumption of selection on observables, one would then conclude that the return to the grant does not come all from the first $25,000 and then fall sharply, but instead there appears to be higher performance gains from granting more funding. If in contrast this reflects selection on unobservables, these results suggest that more funding was going to firms with greater growth prospects.

Appendix Table A16.3: Association between Round 3 outcomes and grant amount received among experimental winners (conditional on receiving all 4 tranche payments).

	Operates Firm	Total Employment	10+ Workers	IHS Profits	Profits and Sales Index
Panel A: New Firms					
no additional controls					
Grant Received (in millions)	0.003	0.635***	0.042***	-0.045	0.022***
	(0.005)	(0.128)	(0.009)	(0.087)	(0.008)
with baseline controls					
Grant Received (in millions)	0.001	0.619***	0.041***	-0.058	0.022**
	(0.006)	(0.138)	(0.009)	(0.094)	(0.009)
Sample Size	375	367	367	365	365
Panel B: Existing Firms					
no additional controls					
Grant Received (in millions)	0.000	0.703***	0.038***	0.137	0.049***
	(0.005)	(0.172)	(0.010)	(0.092)	(0.017)
with baseline controls					
Grant Received (in millions)	0.000	0.678***	0.040***	0.164	0.030*
	(0.006)	(0.184)	(0.012)	(0.110)	(0.016)
Sample Size	244	240	240	241	242

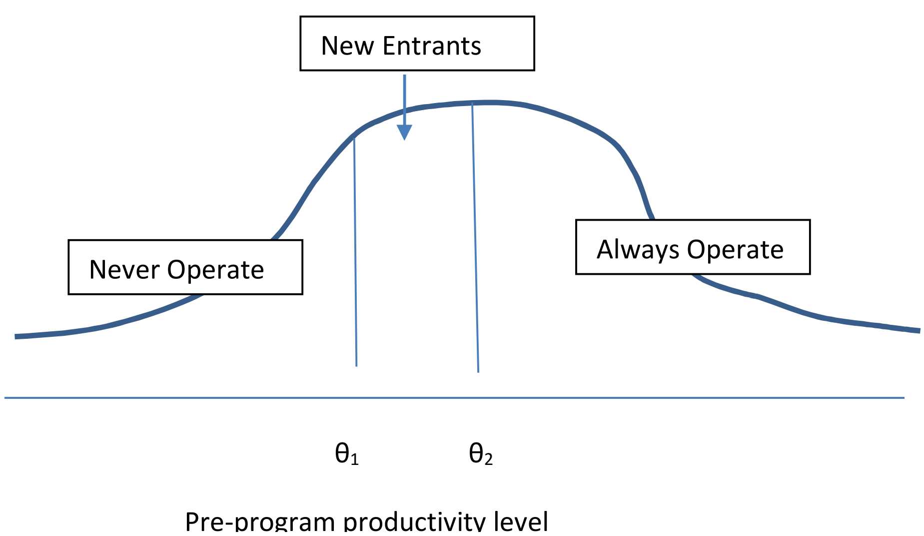
Notes:
Regressions are just on the experimental treatment group, and explore the correlation between round 3 outcomes and the grant amount received (in millions of Naira). Controls are the baseline variables included in Table 1. Robust standard errors in parentheses.
*, **, *** indicate significance at the 10, 5, and 1 percent levels respectively.
New and Existing firms denote status at time of application.
Regressions are for sample receiving all 4 tranche payments.

Appendix 17: Is the Program Resulting in Less Efficient Firms Running Businesses?

Let θ = h(A,E) be the pre-program productivity of a firm owner. Individuals with higher values of θ are able to use their entrepreneurial skills and other productive advantages of their business to generate more output from a given level of inputs K and L. In the Lucas (1978) model, there are no market imperfections, and θ determines the optimal size of the firm, and as discussed in the body of the paper, a grant will not induce any additional business entry.

Consider then the case when finance is limited, but is allocated efficiently so that it goes to the highest productivity individuals. Then there will be a level of θ, say θ_2, such that all individuals with productivity levels of at least θ_2 will operate a business regardless of whether or not they win. In such a model, the additional firms created or induced to survive by winning the competition will have lower productivity than all the control group firms, with the new firms have productivity levels between some cutoff θ_1 and the old cutoff level θ_2. This is illustrated in appendix Figure A17.1

Appendix Figure A17.1: The Possibility that Winning Results in Lower Productivity Firms Entering or Surviving.

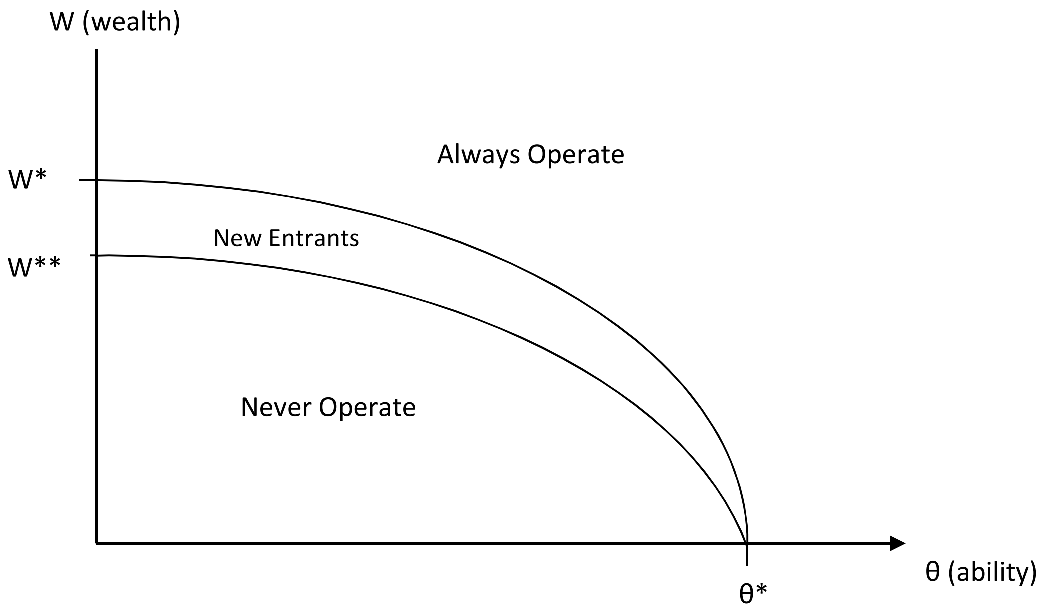


In contrast, if we consider a more general model of occupational choice under credit constraints, such as the model of Lloyd-Ellis and Bernhardt (2000) adapted by de Mel et al. (2014), then business ownership will be jointly determined by a combination of ability and initial wealth (W), as in Figure A17.2.[6] In a simple version of this model, individuals with low wealth and low ability

[6] I am calling this initial wealth for simplicity, but more generally it can be thought of as reflecting how binding credit constraints are, with low wealth corresponding to very limited access to credit.

will never operate a firm, and those with high levels of both will always operate a business. The group induced to operate a firm because of the grant will therefore consist of a mixture of high ability but low wealth individuals, as well as some lower ability but higher wealth individuals. There should therefore be a mix of productivity levels entering, with the specific mix depending on the distribution of talent at different wealth levels.

Appendix Figure A17.2: With liquidity constraints also affecting high ability individuals, it is a more mixed ability group who are induced to enter by winning the program.



If we are in the model in Figure A17.1, there is a high degree of sorting of applicants, and the additional individuals induced to operate a firm will differ substantially in terms of ability and productivity from those who operate firms in the control group. In contrast, in a model like that in Figure A17.2, the additional firms induced to operate will comprise of a much wider mix of ability levels, and we should not see such sorting. It is then an empirical question as to how different the additional firms induced to operate are from those that would operate anyway, and I examine this in several ways

Comparison in terms of baseline characteristics

Appendix Table A17.1 compares characteristics at the time of application for those who are operating a firm at the time of the third follow-up survey in the treatment and control groups. Recall that there are large treatment impacts on the likelihood of operating a firm at the time of the third round survey (37 percentage points for the new firms and 20 percentage points for the existing firms). These large treatment effects result in many more firms being operated in the treatment group than in the control group. However, despite this, we see the baseline characteristics of those operating firms are very similar on average in the treatment and control groups. Those operating firms have similar education levels, a similar mix of industries, and have similar business plan scores. The joint orthogonality tests cannot reject the null hypothesis that these characteristics do not jointly predict treatment status. This evidence suggests that the additional firms induced to operate by winning are not more marginal in terms of these characteristics. In particular, they do not differ significantly in terms of average business plan score.

Comparison in terms of ability and personality

The first follow-up survey administered a Raven test (a measure of non-verbal IQ), a digitspan recall test (a measure of short-term cognitive capacity), and Duckworth et al. (2007)'s 12-item grit measure. Grit measures the extent to which individuals display passion and perseverance towards their goal. Appendix Table A17.2 shows that there is no significant difference on any of these three measures between the treatment and control groups. These results suggest that winning is not just inducing lower ability or less passionate individuals who would not otherwise operate firms to enter, but is drawing from across the distribution in terms of these attributes.

Comparison in terms of realized labor productivity

Measuring productivity is challenging and the subject of much debate in the literature. This is particularly the case for comparing multi-product firms in a wide variety of industries with one another, without good data on input or output prices. I therefore focus on sales and profits in the body of the paper. However, a crude measure of productivity is sales per worker, sometimes referred to as labor productivity. To examine how the additional firms induced to operate compare in terms of labor productivity to the control group firms, I employ the following steps:

1. Determine the quartiles of sales per worker for the control group firms that are operating firms in round 3, separately for new and existing firms, and then count the number of control group firms in each of these quartiles.[7] For example, there are between 84 and 86 new firms in each quartile for the control group.
2. Multiply the number of new firms in the quartile by the ratio of treatment firms to control firms for which we have round 3 data (whether in business or not). This gives the predicted number of treatment firms in that quartile if there had been no treatment effect. For example, since there are 392 treatment firms and 693 control firms with data in round 3 for new firms, we expect there to be (392/693)*84 = 48 treated firms in the same quartile as there were 84 control firms.
3. Then count how many treated firms there actually are in this sales per worker quartile for the treated firms, separately for new and existing firms. For example, there are 83 treated firms amongst the new applicants in round 3 for which the labor productivity is at a level that would place them in the bottom quartile of the control group.
4. Compare the actual numbers from step 3 to the predicted numbers in step 2 to calculate the number of additional firms with that productivity level in the treated group. Then bootstrap

[7] The number is of course almost the same for each quartile, with the difference due to rounding and odd numbers.

to get a confidence interval which allows for the sampling variability in this prediction error.

Figure A17 shows the results of this exercise. For new firms we see the largest share of additional firms comes in the second quartile of the control group productivity distribution, accounting for 48 percent of the additional new entrants. Then slightly more of the remaining additional firms fall into the third quartile of the distribution than the first, with no significant additional firms coming in the top quartile. As a result, the majority of the additional firms operating after winning are more productive than the bottom quartile of the control group firms operating, but we can't rule out that they are all less productive than the top quartile of the control group firms operating.

In contrast, the additional firms induced to operate and survive amongst the existing firms are more evenly spread out across the control group productivity distribution, with 17 percent in the lowest quartile, and 26 to 29 percent in the other three quartiles. In particular, here 83 percent of the additional existing firms are more productive than the lowest quartile of the control group firms in operation, and 29 percent are more productive than three-quarters of the control group firms.

Since this exercise is based on realized productivity, it reflects the combination of selection effects as to which firms enter and survive, and treatment effects (the extent to which winning changes productivity). In either case, the evidence is not consistent with the fear that the competition is only increasing operation of the least-productive, least-talented firms.

Taken together these results are therefore more in line with a model where finance is not allocated to the most efficient firms, and so a broad mix of ability levels are induced to operate firms through the program.

Appendix Table A17.1: Comparison of Baseline Characteristics for Firms Operating in Round 3

	Existing Firms			New Firms		
	Treatment	Control	p-value	Treatment	Control	p-value
Applicant Characteristics						
Female	0.174	0.153	n.a.	0.159	0.123	n.a.
Age	32.2	32.0	0.568	29.5	30.0	0.171
Married	0.517	0.547	0.466	0.341	0.382	0.292
High School or Lower	0.128	0.135	0.888	0.098	0.102	0.417
University education	0.616	0.606	0.857	0.718	0.682	0.063
Postgraduate education	0.070	0.100	0.342	0.039	0.061	0.325
Lived Abroad	0.083	0.100	0.576	0.059	0.067	0.834
Choose Risky Option	0.554	0.506	0.315	0.553	0.521	0.308
Have Internet access at home	0.554	0.600	0.316	0.453	0.484	0.779
Own a Computer	0.888	0.876	0.686	0.841	0.874	0.391
Satelite Dish at home	0.669	0.700	0.520	0.690	0.676	0.810
Freezer at home	0.558	0.618	0.220	0.517	0.556	0.425
Business Characteristics						
Crop and Animal Sector	0.165	0.153	0.681	0.237	0.209	0.487
Manufacturing Sector	0.264	0.271	0.795	0.293	0.238	0.165
Trade Sector	0.058	0.059	0.955	0.034	0.048	0.413
IT Sector	0.149	0.153	0.935	0.064	0.061	0.787
First Round Application Score	57.0	55.9	0.077	60.1	59.9	0.445
Business Plan Score	45.7	44.8	0.212	53.6	55.2	0.945
Number of Workers	6.941	7.946	0.288			
Ever had Formal Loan	0.056	0.083	0.312			
Joint orthogonality test p-value	0.969			0.603		

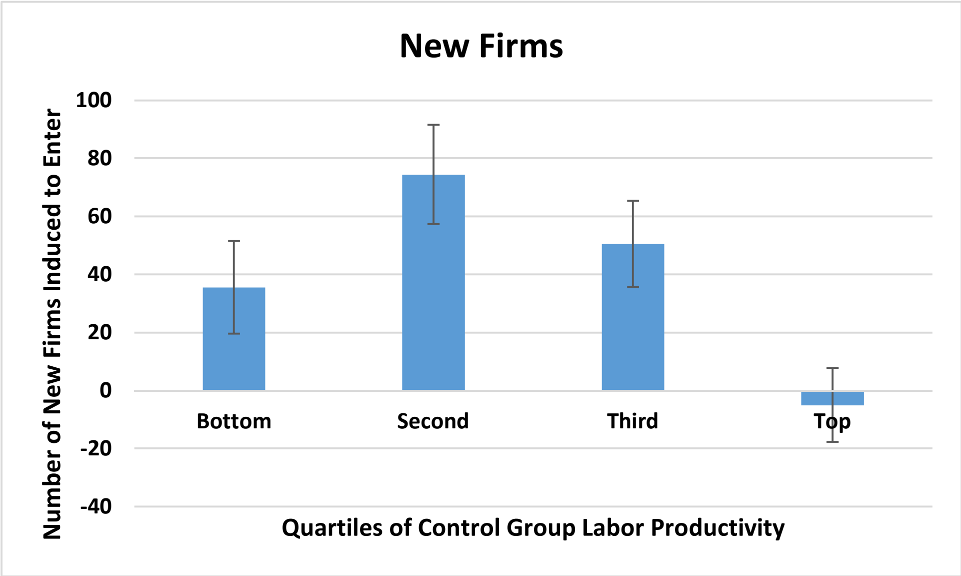
Note: p-values are from regression controlling for randomization strata, so are not available (n.a.) for gender

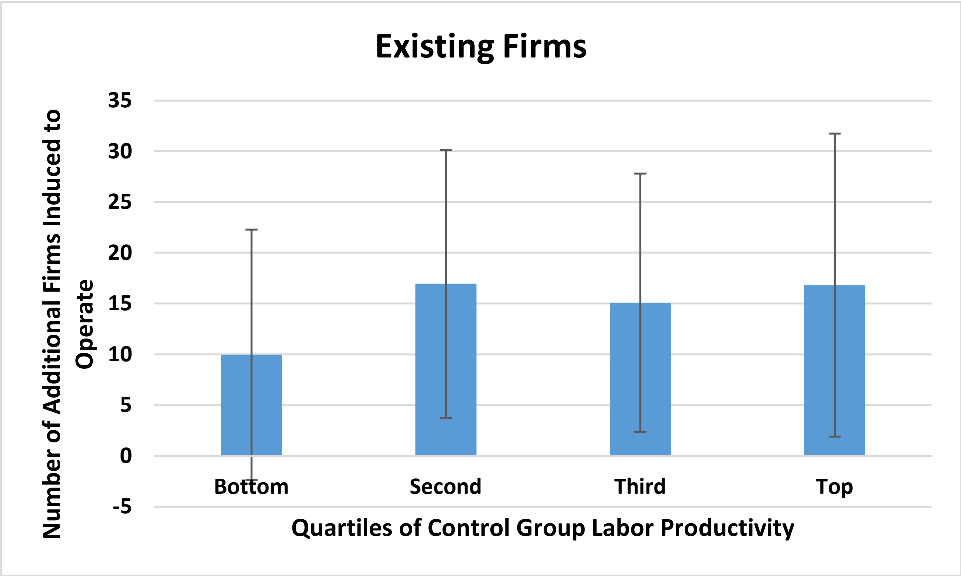
Appendix Table A17.2: Comparison of those operating firms in round 3 in terms of ability and grit

Comparison of Ability and Personality

	Existing Firms			New Firms		
	Treatment	Control	p-value	Treatment	Control	p-value
Raven test score	4.44	4.37	0.889	4.38	4.29	0.359
Digitspan recall	7.39	7.46	0.638	7.26	7.58	0.325
Grit	3.97	4.01	0.594	3.81	3.86	0.903

Appendix Figure A17: Where are the additional treated firms in terms of the control group productivity distribution?





Notes: Figures show how many more treated firms are operating with a given labor productivity level in round 3 than would be predicted based on the productivity distribution of the control group firms in operation. 95 percent confidence intervals that allow for prediction uncertainty shown.

Appendix 18: Heterogeneity in Impacts for Existing Firms

Three key (pre-specified) dimensions of heterogeneity of policy interest for the program are whether impacts are larger for new or existing firm applicants, how impacts vary with gender, and how impacts vary with the business plan score. Although officials believed that supporting existing firms would generate more jobs than new firms, Table 3 and Appendix 11 shows grants to new firms actually generated more jobs per firm supported.

Tables A18.1 and A18.2 then examines the heterogeneity in treatment effect with respect to gender and the business plan score for new and existing firms respectively. Recall that females were only 18 percent of applicants and 17.6 percent of winners. Among the new firm applicants, females in the control group were less likely to start a business during the three years after applying (42 percent operate a business after three years, versus 56 percent of male applicants). The grants help to close this gender gap, having a positive and significant interaction effect with treatment. In contrast, despite female-operated firms having lower employment and sales and profits, treatment does not have significant impact on closing this gap. For existing firms, female-owned existing firms are not less likely to survive than male-owned firms, and there is no significant gender treatment interaction for these firms.

The experimental winners encompass a wide range of business plan scores, from 30 to 73, with a standard deviation of 9. In the first year, the grants had more impact for treated new firm applicants with lower business plan scores, again closing a gap that would otherwise exist. There is a small, but significant, negative interaction of treatment with the business plan score for round 2 profits and sales. But there is no significant heterogeneity with respect to business plan score by the third round, or for existing firms.[8] The results do not support the idea that the program would have had more impact if grants had been given to the semi-finalists with the highest scores, rather than randomly given for firms with scores in this range.

[8] Moreover, the heterogeneity found is not robust to correcting for multiple hypothesis testing. Taking the 18 interactions tested in Table 7, the lowest sharpened q-value is 0.147, for both the round 1 interaction of treatment with gender when business operation is the outcome, as well as the round 1 interaction of treatment with business plan score.

Appendix Table A18.1: Heterogeneity in Impacts for New Firms

	Operates a Firm			Total Employment			Profits and Sales Index		
	Round 1	Round 2	Round 3	Round 1	Round 2	Round 3	Round 1	Round 2	Round 3
Panel A: Heterogeneity by Gender									
Assigned to Treatment	0.185***	0.341***	0.354***	1.411	6.119***	4.976***	-0.002	0.282***	0.159***
	(0.032)	(0.025)	(0.026)	(0.860)	(0.471)	(0.508)	(0.054)	(0.041)	(0.048)
Assigned to Treat*Female	0.189**	0.104*	0.120*	0.100	-0.638	1.557	0.119	0.093	0.055
	(0.078)	(0.063)	(0.067)	(1.021)	(0.895)	(1.321)	(0.084)	(0.077)	(0.099)
Sample Size	1021	1181	1085	987	1159	1044	995	1152	1063
Control Mean Females	0.420	0.481	0.422	1.674	2.165	2.883	-0.233	-0.239	-0.144
Control Mean Males	0.574	0.586	0.562	3.964	3.539	3.937	0.035	-0.067	-0.032
Panel B: Heterogeneity by Business Plan Score									
Assigned to Treatment	0.677***	0.484***	0.371***	-0.305	3.462	2.456	0.451	0.723***	0.211
	(0.173)	(0.141)	(0.141)	(3.642)	(2.274)	(2.686)	(0.319)	(0.210)	(0.241)
Assigned to Treatment*Business Plan Score	-0.009***	-0.002	0.000	0.032	0.047	0.051	-0.008	-0.008**	-0.001
	(0.003)	(0.003)	(0.003)	(0.075)	(0.043)	(0.051)	(0.006)	(0.004)	(0.004)
Business Plan Score	0.004	0.001	0.001	0.053	0.054*	0.071*	0.012**	0.007**	0.005
	(0.003)	(0.003)	(0.003)	(0.037)	(0.032)	(0.039)	(0.005)	(0.003)	(0.004)
Sample Size	1021	1181	1085	987	1159	1044	995	1152	1063
Control Mean Bottom Quartile	0.513	0.548	0.563	3.497	3.263	4.372	-0.064	-0.131	-0.065
Control Mean Top Quartile	0.592	0.604	0.567	3.542	3.556	4.353	0.081	-0.034	-0.062

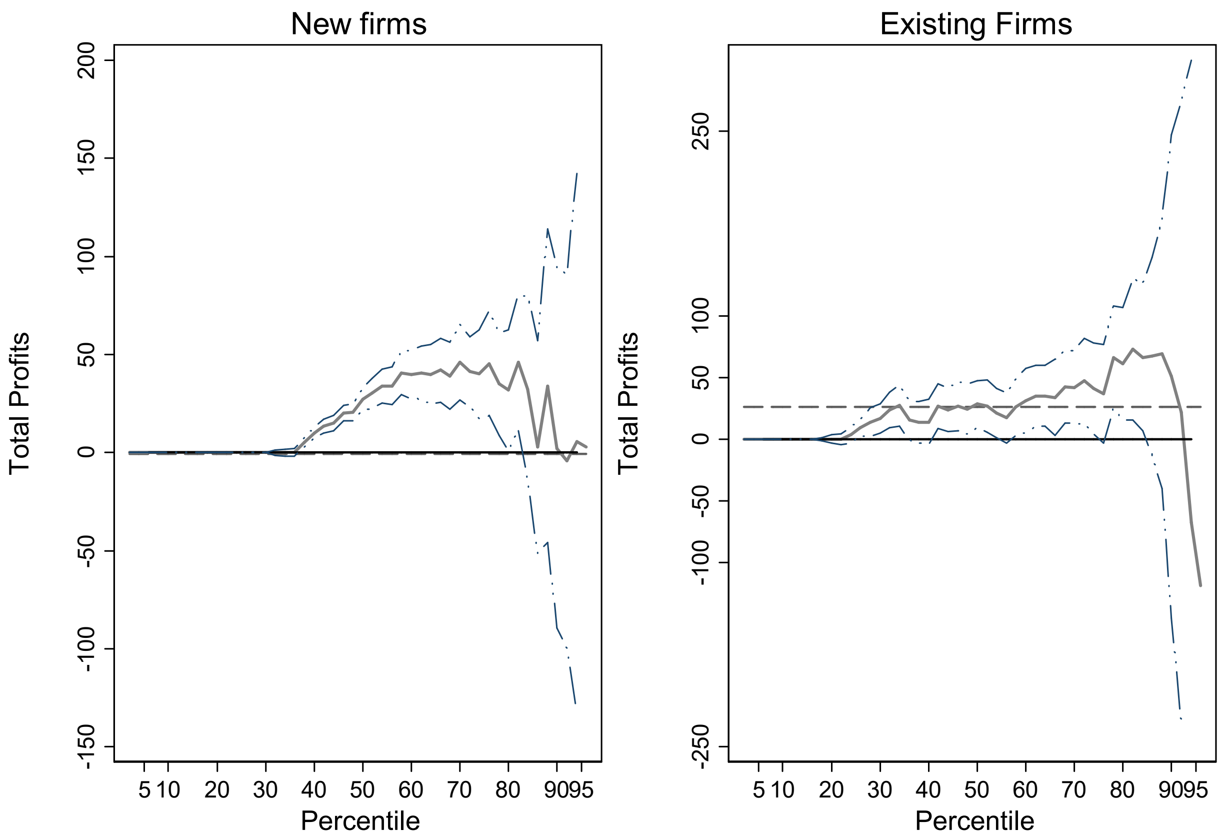
Notes:
Robust standard errors in parentheses, *, **, *** indicate significance at the 10, 5, and 1 percent levels.
Experimental estimates are ITT estimates and control for randomization strata.
Existing and New refers to firm status at time of application. Rounds 1, 2, and 3 are 1, 2, and 3 years after application.

Appendix Table A18.2: Heterogeneity in Impacts for Existing Firms

	Operates a Firm			Total Employment			Profits and Sales Index		
	Round 1	Round 2	Round 3	Round 1	Round 2	Round 3	Round 1	Round 2	Round 3
Panel A: Heterogeneity by Gender									
Assigned to Treatment	0.092***	0.139***	0.185***	1.553*	2.176	4.348***	0.073	0.242***	0.225***
	(0.032)	(0.029)	(0.035)	(0.864)	(1.628)	(0.685)	(0.076)	(0.066)	(0.076)
Assigned to Treat*Female	-0.060	-0.051	0.061	-0.267	2.182	0.477	0.049	-0.033	-0.086
	(0.045)	(0.059)	(0.083)	(2.211)	(2.571)	(2.293)	(0.191)	(0.162)	(0.194)
Sample Size	432	505	477	422	500	461	423	497	470
Control Mean Females	0.967	0.886	0.722	7.862	7.364	6.091	0.027	-0.048	-0.009
Control Mean Males	0.854	0.834	0.766	6.669	8.309	5.475	-0.058	-0.132	-0.126
Panel B: Heterogeneity by Business Plan Score									
Assigned to Treatment	-0.022	-0.041	-0.047	1.153	9.559*	6.290	0.477	0.461	-0.342
	(0.166)	(0.154)	(0.201)	(4.506)	(5.577)	(4.671)	(0.471)	(0.357)	(0.466)
Assigned to Treatment*Business Plan Score	0.002	0.004	0.005	0.008	-0.154	-0.040	-0.009	-0.005	0.012
	(0.004)	(0.003)	(0.004)	(0.101)	(0.140)	(0.099)	(0.011)	(0.008)	(0.011)
Business Plan Score	-0.004	-0.005	-0.004	-0.007	0.112	-0.025	0.003	-0.003	-0.006
	(0.003)	(0.003)	(0.004)	(0.083)	(0.117)	(0.063)	(0.009)	(0.006)	(0.007)
Sample Size	432	505	477	422	500	461	423	497	470
Control Mean Bottom Quartile	0.944	0.844	0.814	7.962	6.391	6.456	0.027	-0.167	0.025
Control Mean Top Quartile	0.852	0.849	0.778	7.698	10.873	5.417	0.072	-0.093	-0.092

Notes:
Robust standard errors in parentheses, *, **, *** indicate significance at the 10, 5, and 1 percent levels.
Experimental estimates are ITT estimates and control for randomization strata.

Appendix 19: Quantile Treatment Effects for Profits in Round 4



Note: Figure shows quantile treatment effect of treatment on monthly profits in round 4, taken 3 years after treatment. Dashed horizontal line shows OLS treatment effect.

Appendix 20: Spillovers

Appendix Table A20.1 shows that the treatment effect does not vary with the number of other winners in the same state and industry. 16 percent of the control firms have no winners in their state and industry, 27 percent have 1-2 firms, 18 percent have 3-4 firms, and 35 percent have 5 or more. We see that the control group outcomes are no different when they face more winners, and that the treatment effect is not different for firms facing more winners.

Appendix Table A20.1: Heterogeneity in Treatment Effect With Respect to Number of Other Winners in the Same State and Industry

	Outcomes in Third Follow-up					
	New Firms			Existing Firms		
	Operates a Firm	10+ workers	Sales & Profits Index	Operates a Firm	10+ workers	Sales & Profits Index
Assignment to Treatment	0.388***	0.230***	0.224***	0.202***	0.256***	0.264***
	(0.033)	(0.037)	(0.057)	(0.046)	(0.060)	(0.092)
assigntreat_numberotherwinners	-0.004	-0.000	-0.013	-0.002	-0.010	-0.010
	(0.005)	(0.005)	(0.009)	(0.007)	(0.009)	(0.012)
numberotherwinners	-0.003	0.000	0.001	-0.001	0.002	0.010
	(0.004)	(0.003)	(0.007)	(0.006)	(0.006)	(0.011)
Sample Size	1085	1044	1063	477	461	470

Notes:
All regressions also include controls for randomization strata. Robust standard errors in parentheses.
*, **, *** indicate significance at the 10, 5, and 1 percent levels respectively.

Most firms in the study produce and sell a wide variety of products, making it difficult to measure changes in their prices. The first and second round surveys asked firms the price per unit of their main product. I use this to examine whether treated firms were more likely to lower prices between rounds 1 and 2, during the period in which they expanded their firm size and received most of the grant. This outcome was not included in the pre-analysis plan, and conditions on survival, so should be considered suggestive only. It shows that winners were no more likely to have lowered the price of their main product than the control group.

Appendix Table A20.2: Treated firms are not more likely to have lowered prices than control firms

Dependent variable: lowered price between round 1 and round 2		
	Existing Firms	New Firms
Assignment to Treatment	0.005	-0.032
	(0.055)	(0.047)
Sample Size	341	469
Control Mean	0.436	0.517

Notes:
Price is price per unit of main product sold. Data only for sample in business in round 1 and round 2, and so conditions on start-up and survival.
Robust standard errors in parentheses, *, **, *** indicate significance at the 10, 5, and 1 percent levels respectively.

References cited only in the Appendices:

Anderson, Michael (2008), "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects", *Journal of the American Statistical Association*, 103(484), 1481-1495

Attanasio, Orazio, Adriana Kugler, and Costas Meghir, (2011). "Subsidizing vocational training for disadvantaged youth in Colombia: Evidence from a randomized trial", *American Economic Journal: Applied Economics* 3(3): 188-220

Behaghel, Luc, Bruno Crépon, Marc Gurgand, Thomas Le Barbanchon (2015) "Please Call Again: Correcting Non-Response Bias in Treatment Effect Models", *Review of Economics and Statistics*, 97(5): 1070-1080

Benjamini, Yoav, Abba M. Krieger, and Daniel Yekutieli (2006) "Adaptive Linear Step-Up Procedures That Control the False Discovery Rate." *Biometrika* 93 (3): 491–507.

Blattman, Christopher and Laura Ralston (2015) "Generating employment in poor and fragile states: Evidence from labor market and entrepreneurship programs", Mimeo. Columbia University.

Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts (2013) "Does management matter? Evidence from India", *Quarterly Journal of Economics*, 128(1): 1-51

Bruhn, Miriam, Dean Karlan and Antoinette Schoar (forthcoming) "The Impact of Consulting Services on Small and Medium Enterprises: Evidence from a Randomized Trial in Mexico", *Journal of Political Economy.*

Bruhn, Miriam and David McKenzie (2014) "Entry regulation and formalization of microenterprises in developing countries", *World Bank Research Observer*, 29(2): 186-201.

De Mel, Suresh, David McKenzie and Christopher Woodruff (2012) "One-Time Transfers of Cash or Capital Have Long-Lasting Effects on Microenterprises in Sri Lanka", *Science* 335: 962-966.

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

Drexler, Alejandro, Greg Fischer, and Antoinette Schoar (2014) "Keeping it Simple: Financial Literacy and Rules of Thumb", *American Economic Journal: Applied Economics* 6(2): 1-31.

Duckworth, A.L., Peterson, C., Matthews, M.D., & Kelly, D.R. (2007) "Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*", 9: 1087-1101.

Groh, Matthew, Nandini Krishnan, David McKenzie and Tara Vishwanath (2016) "Do Wage Subsidies Provide a Stepping Stone to Employment for Recent College Graduates? Evidence from a Randomized Experiment in Jordan", *Review of Economics and Statistics*, 98(3): 488-502.

Grimm, Michael and Anna-Luisa Paffhausen, (2015) "Do Interventions Targeted at Micro-entrepreneurs and Small and Medium-sized Firms Create Jobs? A Systematic Review of the Evidence for Low and Middle Income Countries", *Labour Economics*, 32: 67-85.

Guerrieri, Veronica and Daron Acemoglu (2008) "Capital Deepening and Non-balanced economic growth", *Journal of Political Economy* 116(3): 467-98.

Hirschleifer, Sarojini, David McKenzie, Rita Almeida and Cristobal Ridao-Cano (2016) "The Impact of Vocational Training for the Unemployed: Experimental Evidence from Turkey" *Economic Journal*, 126(597): 2115-46.

Horowitz, Joel L. and Charles F. Manski (2000) “Nonparametric Analysis of Randomized Experiments with Missing Covariate and Outcome Data”, *Journal of the American Statistical Association* 95(449): 77-84.

Imbens, Guido and Charles Manski (2004) “Confidence intervals for partially identified parameters”, *Econometrica* 72: 1845-1857.

Karlan, Dean and Martin Valdivia (2011) “Teaching entrepreneurship: Impact of business training on microfinance clients and institutions”, *Review of Economics and Statistics* 93(2): 510-27.

Lee, David (2009) “Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects", *Review of Economic Studies* 76(3): 1071-1102.

Lloyd-Ellis, Huw and Dan Bernhardt (2000) “Enterprise, inequality, and economic development”, *Review of Economic Studies* 67: 147-68.

McKenzie, David (2012) “Beyond Baseline and Follow-up: The Case for More T in Experiments”, *Journal of Development Economics* 99(2): 210-21.

McKenzie, David and Christopher Woodruff (2015) “Business Practices in Small Firms in Developing Countries’, *Management Science*, forthcoming

Valdivia, Martin (2015) “Business training plus for female entrepreneurship? Short and medium-term experimental evidence from Peru”, *Journal of Development Economics* 113: 33-51.