Business Collaborations and Female Entrepreneurship Pre-Analysis Plan

Edward Asiedu Monica Lambon-Quayefio Francesca Truffa
Ashley Wong

1 Introduction

In this study, we run a field experiment with 1,772 female entrepreneurs in Ghana to investigate how a matching service and legal support can affect interfirm collaborations and improve performance of female-owned businesses.

This documents serves as our pre-analysis plan, which outlines the research design, key hypotheses and empirical strategy of the study. In Section 2, we provide an overview of the study. In Section 3, we describe our primary, secondary and immediate outcomes of interest. Finally, in Section 4, we describe the estimation strategy and key hypotheses.

2 Overview of the Study

2.1 Motivation and Program Description

In developing countries, female-owned businesses are smaller and earn 34% less profits than male-owned firms (World Bank Group, 2019). One potential explanation is that female entrepreneurs have smaller networks and are less connected to other firms (World Bank Group, 2019). Importantly, studies suggest that forming interfirm relationships such as strategic partnerships and collaborations can be important for business growth and innovation by allowing firms acquire new customers, expand market reach, reduce costs, gain access to new technologies and intellectual property (Kanter, 1994). However, it remains

an open question in the literature on the determinants of collaborations and how interfirm relationships contribute to the growth of female-owned businesses.

In this study, we explore whether providing a matching service combined with legal support can increase firm collaborations and improve business outcomes. Specifically, we will provide the following two treatments. Their effect will be compared to the status of a control group receiving no support:

Treatment 1 – Matching Service This treatment aims to reduce the *networking frictions* for collaborations. The entrepreneurs in this treatment arm will be assigned to WhatsApp groups of 8 entrepreneurs which are matched by preferences for collaboration type and sector. All women will participate in weekly virtual "coffee chats", i.e. one-on-one meetings between group members. After everyone in the group has met each other ($\approx$ 8 weeks, 1 for the group introduction and 7 for the one-on-one meetings), they will be re-assigned into a second group and the same process repeated, for a total of approximately 16 weeks. Additionally, the treatment group will also receive access to an online directory of businesses in the respective treatment group and can submit specific requests for collaboration partners to the enumerators who will help connect them to another firm in the sample.

Treatment 2 – Matching Service + Legal Support This treatment arm aims to reduce not only search frictions, but also *contracting frictions* for collaboration. In addition to the support described in Treatment 1, entrepreneurs in this group will also receive legal support. Specifically, they will receive weekly video lessons by a Ghanaian corporate lawyer. These lessons will focus on benefits and risks of collaborations as well as on how to mitigate these risks. In Appendix A, we present the course syllabus. Finally, the entrepreneurs will also receive free private consultations with the lawyer who will be available for phone calls during weekly "office hours" throughout the four-month intervention period.

Finally, given that our treatments are explicitly aiming at increasing collaboration, part of

our effect may be driven by a differential effect of collaboration salience between treated and control groups. In order to mitigate this concern, the entire sample of participants including the control group will be provided with a video illustrating the benefits of collaboration.

2.2 Sample Selection and Assignment to Treatment

Our sample comes from the applicant pool of the COVID-19 Stimulus Fund, offered by our partners Women's Empowerment & Investment Group (WEIG), Annan Capital Partners (ACP), and GUBA Foundation. Specifically, the COVID-19 Stimulus Fund offered funding of $2,000-$5,000 to female-owned small and medium enterprises. The total 3,931 applicant firms form the main sampling frame for the study. The 10 firms that were selected to receive funding were dropped from the research study. We randomly selected 2,326 firms from the pool of applicants who provided their email addresses and phone numbers, and answered all application questions to be included as part of the baseline survey. We then applied an eligibility filter to determine who can be part of the survey: firms with at least one female owner, at least 18 years old, have started business operations at the time of the survey, have at least one business, can speak English or Twi, and provided information on the firm industry and region. Baseline survey was conducted between October 2020 and December 2020 for 2,000 firms. In December 2020, enumerators conducted a short phone survey to elicit interest in the WhatsApp groups and preferences for collaborations. Out of the 2,000 firms, 1,488 (74%) indicated interest and had WhatsApp capability to participate in the study. In January 2021, we contacted an additional 326 entrepreneurs, among whom 284 were interested in the matching service. We also conducted the baseline survey for this additional sample.

This resulted in a final sample of 1,772. These individuals were then randomly assigned into the two treatment groups and one control group:

- Treatment 1: Matching 40% ($N = 704$)
- Treatment 2: Matching + Legal Support 35% ($N = 608$)
- Control 25% ($N = 436$)

We stratify the randomization based on above and below median of the predicted collaboration index, 4 broad sectors ("Crop and animal", "Manufacturing", "Trade", and "Services"), and 5 broad regions ("Ashanti", "Eastern", "Northern", "Volta", and "Western"). The predicted collaboration index is constructed by predicting the likelihood of having at least one collaboration in the past 6 months using random forest. In Table 1, we present the fifteen most important predictors of collaborations selected by the random forest.[1] The reason why we stratify over the predicted collaboration index as opposed to the baseline value of collaboration is because we only collected collaboration information for a subsample of 904 firms.

2.2.1 Assignment to WhatsApp Groups

As part of the intervention, we assign individuals into WhatsApp groups of 8 based on their preferences over collaboration. We elicited preferences over their top three choices for collaboration types, preferred industry of the potential collaborator, and preferred location of the potential collaborator. We then use a two-stage procedure to assign individuals into networking groups. First, within treatment status and language (English or Twi), we assign individuals into one of 25 group types based on their preferences over 5 collaboration types (creation of new products with collaboration, joint marketing, joint production, finding suppliers/clients, and mixed types[2]) and five broad sector group (crop and animal, manu-

[1] More details on the machine learning prediction are provided in Appendix Section C.

[2] The mixed group comprise of any remaining individuals for whom we could not group based on their collaboration preference.

facturing, trade, services, and mixed sector).[3] Then, in the second stage of the assignment, individuals are randomly selected to be placed into groups of 8 within their group type. To help identify the importance of group composition, we randomly select half of the sample to be placed in a group with individuals with the same education background (either college-educated or not) and the other half are placed into mixed education groups. The purpose is to generate variation along the education dimension across groups to help us identify key mechanisms.

2.3 Data collection

Our baseline survey took place between October and December 2020. We collected information on key firm and owner characteristics. For a subsample of around 900 firms we also collected detailed information on collaboration. In December 2020, we conducted a short phone survey to collect information on interest in the matching program and collaboration preferences in terms of collaboration type, sector, and location of the potential partner. In June and December 2021, we plan to conduct our midline and endline surveys, respectively. Moreover, throughout the duration of our project (from October 2020 to December 2021), we will collect real-time business transaction data through OZÉ, a Ghana-based business management mobile phone application. Specifically, all entrepreneurs in our sample were provided access to a 3-month premium subscription to OZÉ through our NGO partner. This application helps businesses record transactions and send receipts. Through our partnership with OZÉ, will collect real-time business-level data on sales, expenses, profits, and cash flow aggregated from transactions.[4] We will use this data to supplement the data we collect through our surveys.

[3]Due to the limited sample size, we decided not to account for preferences over location in our group assignment. Additional details on the assignment process for the group types are provided in Appendix B.

[4]Note that we will only have this data for the sample of entrepreneurs who choose to use this mobile phone application.

3 Outcomes

We split our outcomes into primary, secondary, and intermediate outcomes. The details for each outcome are provided in the appendix.

3.1 Primary Outcomes

Hypothesis 1: Both matching and matching with legal support will increase collaborations

Summary measures:

- Number of Collaborations
- Steps towards Collaboration Index
- Joint application for business innovation competition

In Appendix Section D.1, we describe how we define these outcome measures and explain the business innovation competition.

3.2 Secondary Outcomes

Our secondary outcomes consist of outcomes that we are interested in investigating, but we are likely under-powered to detect.

Hypothesis 2a: Both matching and matching with legal support may lead to firm innovation.

Summary measure: Business Innovation index

Hypothesis 2b: Both matching and matching with legal support may improve firm performance.

Summary measure: Sales and profits index

Hypothesis 2c: Both matching and matching with legal support may increase the quality of collaborations.

Summary measure: Collaboration Quality Index

Details on how the summary measures for the secondary outcomes are defined can be found in Appendix Section D.2.

3.3 Intermediate Outcomes

Hypothesis 3a: Both matching and matching with legal support may improve additional business outcomes.

Summary measure:

- Business practices index
- Business financing index
- Capital and labor index

Details on the summary measure variables can be found in Appendix Section D.3.1.

Hypothesis 3b: Both matching and matching with legal support may improve business networking.

Summary measure: Business networking index

We describe the index construction in Appendix Section D.3.2.

Hypothesis 3c: Matching with legal support may improve firm formalization, collaboration knowledge and trust in the legal system.

Summary measure:

- Use of legal and consulting services
- Use of written agreement

- Business formalization index
- Knowledge related to collaborations
- Trust in contracts
- Trust in the legal system

Details on the construction of the business formalization index as well as on our measures of trust can be found in Appendix Section D.3.3.

Hypothesis 3d: Both matching and legal support may improve business ambition, attitudes, entrepreneurial self-efficacy, and empowerment.

Summary measure:

- Business Ambitions Index
- Get Ahead Attitudes (from McKenzie and Puerto, 2020)
- Entrepreneurial Self-Efficacy
- Female Empowerment Index

We describe the index construction in Appendix Section D.3.4.

Hypothesis 3e: Both matching and matching with legal support may increase perceived benefits of collaborations.

Summary measure:

- Importance of collaboration to firm growth: self-reported rating on a scale of 0-10, 10 being the most important

Hypothesis 3f: Matching with legal support may increase trust among collaborators more than matching service alone.

Summary measure:

- Collaboration Trust index

We describe the index construction in Appendix Section D.3.5.

Hypothesis 3g: Matching and matching with legal support may increase knowledge about collaboration through peer effects.

We will test whether collaboration is more likely when peers are more likely to have collaborated in the past or if peers are more educated. We describe this test further in Section 4.1.1

Hypothesis 3h: Matching and matching with legal support may change the types of firms with whom firms collaborate.

Summary measure:

- Probability of collaborating with a friend or family member
- Probability of collaborating with a firm met through a business network

Hypothesis 3i: Matching with legal support may lead to more formal collaborations than matching service alone.

Summary measure:

- Collaboration Formalization index

We describe the index construction in Appendix Section D.3.6.

4 Analysis

4.1 Estimation Methodology

To investigate the effects of our treatments on our outcomes of interest, we will estimate:

$$Y_{i,t=1} = \beta_0 + \beta_1 T1_i + \beta_2 T2_i + \pi Y_{i,t=0} + \delta M_{i,t=0} + S_i'\gamma + \tau \hat{f}(X) + \epsilon_{i,t=1} \quad (1)$$

where β_1 represents the effect of matching service only and β_2 represents the effect of matching service and legal support. $T1_i$ and $T2_i$ are indicators for treatment 1 and 2 , respectively. S_i is the vector of randomization strata dummies. $Y_{i,t=0}$ is the baseline value of the outcome Y. $M_{i,t=0}$ is an indicator if the baseline outcome value was missing at baseline, and $\hat{f}(X)$ is the ML index that predicts Y from controls (Ludwig et al., 2019; Wager et al., 2016; Bloniarz et al., 2016; Wu and Gagnon-Bartsch, 2018).[5] Standard errors are clustered at the WhatsApp group level.

4.1.1 Peer Effects Estimation

In addition to the main specification, we will also explore the role of peers in explaining our results. To do so, we estimate the following linear-in-means model for individuals in the treated groups only:

$$\begin{aligned} Y_{i,t=1} = \alpha_0 + \alpha_1 T2 + \alpha_2 \overline{X}_{-i,t=0} + \alpha_3 \overline{X}_{-i,t=0} \times T2 \\ + \pi Y_{i,t=0} + \delta M_{i,t=0} + S_i'\gamma + \tau \hat{f}(X) + K_i'\phi + \epsilon_{i,t=1} \end{aligned} \tag{2}$$

where $\overline{X}_{-i,t=0}$ is the average characteristic of the peers of i. We will specifically analyze average education level and average baseline predicted collaboration index using ML of peers. In addition to the controls in the main specification, we will additionally control for K_i, a vector of variables used in the group assignment. This includes indicators for treatment status, top collaboration choices, collaboration language preference, interest in collaborating with the same or different sector, firm sector, and all of their interactions. Because the assignment to WhatsApp groups is random conditional on these preferences, α_2 identifies the causal impact of peer composition on our outcome variable and α_3 identifies any difference in peer effects across the two treatment arms.

[5] More details on the machine learning prediction are provided in Appendix Section C.

4.2 Hypothesis Testing

For each outcomes listed, we will test the following hypothesis:

- whether matching service alone has any effect (H0: $\beta_1 = 0$)
- whether matching service and legal support has any effect (H0: $\beta_2 = 0$)
- whether these differ (H0: $\beta_1 = \beta_2$)

Multiple Test Correction

Because in our study we consider multiple primary outcomes, we will adjust adjust for multiple hypothesis testing to minimize the false non-discovery rate (FNR) following Benjamini and Hochberg (1995) and Anderson (2008). Sharpened q-values will be presented by each outcome grouping.

4.3 Index Construction

For some of our outcome variables, we group several related variables into index variables. We construct the indices in four steps. First, we re-code all contributing outcomes so that higher values correspond to treatment effects in the same direction (improvements in the outcomes). Second, we generate z-scores for each variable entering the index using the baseline mean and standard deviation for that outcome. Third, we generate means of these z-scores. Fourth, we create the index by generating the z-score for the means of these z-scores.

4.4 Heterogeneous Effects

We will test heterogeneity of impacts along the following dimensions:

1. An indicator equal to one if respondent has below median baseline business networking index

2. An indicator equal to one if respondent has above median risk aversion measured at baseline

3. An indicator equal to one if respondent has below median baseline business formalization index

4. An indicator equal to one if respondent has below median baseline legal knowledge index

5. An indicator equal to one if respondent has above median education level

4.5 Missing data from item non-response or survey attrition

We will dummy out any missing baseline values in our main specification. For missing data from the follow up surveys, we will not impute and present the results without adjustment for attrition or item non-response. For any outcome that is missing for at least 10% of the sample, we will analyze whether item non-response is correlated with treatment status. If it is then we will construct Lee bounds to account for this in our treatment estimates.

4.6 Questions with Limited Variation

Following McKenzie and Puerto (2020), we will not exclude from construction of any aggregate index questions for which 95 percent of observations have the same value within the sample in order to limit noise.

5 Tables

Table 1: Predictors of Any Collaborations in the Last Six Months using Random Forest

Ten Most Important Predictors
1. Get Ahead Attitude Index
2. Business Practice Index
3. Networking Index
4. Trust Index
5. Innovation Index
6. Business Formalization Index
7. Risk Aversion Index
8. Ambition Index
9. Owner's Age
10. Legal Knowledge Index
11. Self-Efficacy Index
12. Covid-Impact Index
13. Empowerment Index
14. Sales and Profits Index
15. Firm's age

References

Anderson, M. (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.

Benjamini, Y. and Y. Hochberg (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological) 57*(1), 289–300.

Bloniarz, A., H. Liu, C.-H. Zhang, J. S. Sekhon, and B. Yu (2016). Lasso adjustments of treatment effect estimates in randomized experiments. *Proceedings of the National Academy of Sciences 113*(27), 7383–7390.

Cai, J. and A. Szeidl (2017). Interfirm relationships and business performance. *The Quarterly Journal of Economics 133*(3), 1229–1282.

Daugherty, P. J., R. G. Richey, A. S. Roath, S. Min, H. Chen, A. D. Arndt, and S. E. Genchev (2006). Is collaboration paying off for firms? *Business Horizons 49*(1), 61 – 70.

Kanter, R. M. (1994). Collaborative advantage: The art of alliances. *Harvard Business Review*.

Ludwig, J., S. Mullainathan, and J. Spiess (2019). Augmenting pre-analysis plans with machine learning. *AEA Papers and Proceedings 109*, 71–76.

McKenzie, D. (2017, August). Identifying and spurring high-growth entrepreneurship: Experimental evidence from a business plan competition. *American Economic Review 107*(8), 2278–2307.

McKenzie, D. and S. Puerto (2020). Growing markets through business training for female entrepreneurs: A market-level randomized experiment in kenya. *Working Paper*.

Nyaga, G. N., J. M. Whipple, and D. F. Lynch (2010). Examining supply chain relationships: Do buyer and supplier perspectives on collaborative relationships differ? *Journal of Operations Management 28*(2), 101 – 114.

Panayides, P. M. and Y. Venus Lun (2009). The impact of trust on innovativeness and supply chain performance. *International Journal of Production Economics 122*(1), 35 – 46. Transport Logistics and Physical Distribution Interlocking of Information Systems for International Supply and Demand Chains Management ICPR19.

Seppänen, R., K. Blomqvist, and S. Sundqvist (2007). Measuring inter-organizational trust—a critical review of the empirical research in 1990–2003. *Industrial Marketing Management 36*(2), 249 – 265. Project Marketing and the Marketing of Solutions.

Wager, S., W. Du, J. Taylor, and R. J. Tibshirani (2016). High-dimensional regression adjustments in randomized experiments. *Proceedings of the National Academy of Sciences 113*(45), 12673–12678.

World Bank Group (2019). Profiting from parity : Unlocking the potential of women's business in africa. *World Bank*.

Wu, E. and J. A. Gagnon-Bartsch (2018). The loop estimator: Adjusting for covariates in randomized experiments. *Evaluation Review 42*(4), 458–488.

Appendix

A Legal Training Syllabus

1. Benefits of Collaboration (Provided to all treatment and control groups)
    - To explain and discuss what collaboration means and its benefits
2. Types of Collaborations
    - To help participants identify some types of collaboration and identify which type they find feasible for their business
3. Effective Collaborative Activities
    - Information sharing, joint relationship effort, dedicated investment
    - Use of communication as a tool to monitor and prevent risks in collaboration
4. Preparing to collaborate or partner with other firms
    - This will focus on preparatory steps and process for choosing a partner
    - Questions to ask:
        - How does the partnership fit into your bigger strategic picture?
        - How well do you really know your partner?
        - How well have you defined key performance indicators (KPIs) for the partnership?
        - How much analysis and evaluation need to be done keeping your brand and values in mind?
        - Which regulatory matters may affect the negotiations or collaboration and what are the levels of compliance by the intending partners?

5. Protecting your interest prior to collaboration

    - This will cover the processes for securing business interests prior to collaboration
        - Conduct a pre-transaction due diligence (legal, financial operational)
        - Signing a non-disclosure agreement
        - Trademark and intellectual property registration
        - Letter of intent or memorandum of understanding
        - Business registration or incorporation

6. Identifying risks in collaboration

    - To assist participants to identify risk factors for collaboration
    - What risks are associated with collaboration?
        - Performance risk
        - Relational risk
        - Clash of cultures
        - Lack of commitments among management teams
        - Lack of trust
    - Identifying risks by asking the following questions:
        - What risks might be involved in setting up and starting the collaborative partnership?
        - What risks might be involved in organizing the partnership?
        - What risks might be involved in meeting the timelines for the collaboration?
        - What risks might be involved in meeting the objectives of the collaboration?
        - What risks might be involved in not having all the resources/funding needed to manage the partnership?

7. Identifying other common issues with collaboration

    - To identify and discuss other issues that commonly affect the success of any collaboration
        - Inability to reconcile competing interests to attain a union of purpose
        - business entity principle – separating business finances from the individual owner
        - governance and decision-making process
        - share of profits and costs
        - rights, obligations and limitations of the powers of the partners
        - sourcing finance and insurance

8. Documentation and agreements for collaboration

    - To discuss the relevant documentations and agreements for collaboration and reasons for written agreements as a risk control mechanism

9. Dispute resolution

    - To discuss how disputes arise in collaborations and why the process for resolving disputes must be clarified

10. Termination of the collaboration and rights upon termination

    - To discuss the processes and conditions for termination of the collaboration and rights upon termination

11. Promoting Trust in collaborations

    - To identify and discuss ways of promoting trust in collaborations

- What is the value of trust in a collaboration?
- Why is there a lack of trust?
  - Lack of operationalization of processes across boundaries
  - Misalignment in goals KPIs
  - Lack of visibility predictability in pipeline and revenue
- How to achieve trust in a collaboration?
  - Open and effective communication
  - Confidence and predictability
  - New opportunities and exponential gains

B Assignment to WhatsApp Groups

To assign individuals into one of the 25 group types (5 collaboration types $\times$ 5 sector types), we conduct the following procedure:

1. First, we assign everyone to their first choice for collaboration type and sector type. Because we have an excess of firms looking for clients compared to firms looking for suppliers by nearly six times, we randomly allocate 60% of the firms looking for clients to be assigned to their second collaboration choice. Individuals who preferred partners in the same broad sector as theirs are assigned to their specific broad sector group. For those who were interested in partners in another broad sector, we assigned them to the mixed sector group.

2. Once the initial assignment is completed, we create groups of 8 within each group type and treatment status.

   - For supplier and client groups, we create groups of 8 with 4 suppliers and 4 clients,

matched on sector preference. Any suppliers or clients that cannot be matched are reassigned to their next preferred collaboration type.

3. For any remaining unmatched individuals, we match on only their collaboration type preference by re-assigning them to the mixed sector groups. We then form groups of 8 within their group type and treatment status.

4. Finally, we assign the remaining unmatched individuals to the "Mixed" collaboration type and form all remaining groups.

In sum, we construct the groups using the following variables: top three choices of collaboration type, interest in collaborating with the same or different sector, and firm sector.

C Machine Learning Predictions

Our machine learning predictions are constructed using random forest. Specifically, we fit a random forest on our outcomes of interest to obtain the predicted value as well as the most important predictors for all outcome variables. To perform this analysis, we used the R package "h2o". We provided the h2o.randomforest function with the following arguments: the training data frame, the predictor variables (see the full list below), the response variable,[6], the number of trees and maximum tree depth (both chosen with cross-validation).

The list of predictors is the following: any collaboration, sales and profits index, financing index, capital and labor index, business practices index, innovation index, self efficacy index, empowerment index, networking index, formalization index, legal knowledge index, ambitions index, get ahead attitude index, COVID impact index, risk index, trust index, firm's age, age, years of education, married indicator, number of children, an indicator for having a child below 18, an indicator for being a only-women owned firm.

[6]Note that in some of the predictor variables are also outcome variables. When they enter as predictor variables in our model, they are removed from the list of outcome variables.

We replaced missing observations with the value 99 and, for each predictor, we added to the previous list one indicator for whether the predictor is missing.

D Outcome Measures

D.1 Primary Outcomes

Number of Collaborations

We measure total number of collaborations based on the total number of times a firm has engaged in one of the following activities:

- Work with another firm to promote/market each others' businesses or products
- Build a new ongoing working relationship with suppliers or business clients
- Purchase inputs or stocks wholesale with another firm
- Share tools, inputs, equipment or employees with another firm
- Work with another firm to fill a large order
- Start operating business together/sharing of profits with another firm
- Other forms of collaboration

Steps towards Collaboration Index

The index is constructed based on the following measures:

- Identified an area of improvement for your business that may benefit from collaboration with another business
- Considered a collaboration

- Conducted a search process (e.g. asking business network, personal connections) to identify potential collaborators
  - For the treatment group, this includes speaking with enumerators with specific requests for collaborators
- Contacted a specific firm with a proposal to collaborate
- Having multiple conversations oriented towards a collaboration
- Started a collaboration

Joint Application for Business Innovation Competition

As part of the intervention design, we will host a business innovation competition. The competition seeks to fund an innovative business project. The winning firm will be awarded 6,000 GHS. The competition allows for joint applications with one other firm. Joint applications will be awarded 12,000 GHS. This competition will be open to all firms in the sample, including the controls and it will be announced at the beginning of the intervention. We will measure joint applications as a key outcome variable for firm collaborations.

D.2 Secondary Outcomes

Business Innovation Index

We construct this index using the measures listed below following the definition described in McKenzie (2017):

- Introduced new products or service
- Significantly improved an existing product or service previously sold by the firm

- Introduced new or improved business processes (examples might include a new production method, a new quality control system, a new accounting system, or a new delivery system).
- Implemented new design or packaging to give a product a new or significantly changed look, or significantly changed the way you display merchandise.
- Introduced a new channel for selling your goods and services, such as licensing to others, selling in a new type of place, etc.
- Introduced a new method of pricing your goods or services, such as a new type of special offer, or a new way of varying the price according to demand.
- Introduced a new way of promoting or advertising your products or services.
- Changed the way work is organized in your firm, by changing the number of levels in your hierarchy, or the way workers work together, or giving more control over certain processes to other workers in your firm.
- Introduced new quality control standards for suppliers or subcontractors
- Licensed a new technology in the last six months

Sales and Profits Index

We construct this index using the measures listed below following the definition described in McKenzie (2017):

- Monthly sales: Sales in the last month, 0 if not in business
- Monthly sales (Winsorized): Sales in the last month, 0 if not in business. It is winsorized at the 99th percentile.

- Inverse hyperbolic sine of monthly sales: Sales in the last month, transformed
- Profits: total profits of the business in the last month
- Profits (Winsorized): total profits of the business in the last month, winsorized at the 1st and 99th percentile
- Inverse hyperbolic sine of profits: Transformation of profits to allow for zero and negative values of profits
- Profits in the best month: total profits of the business in the best month of the past 12 months
- Profits in the best month (Winsorized): total profits of the business in the best month of the past 12 months, winsorized at the 1st and 99th percentile
- Weekly customers: number of customers the firm has in the past week, winsorized at the 99th percentile. It is zero for firms that are not operating

Quality of Collaborations Index

We construct this index using the following measures. For firms without any collaborations, we impute 0 to each of the measure before constructing the index.

- Quality of the competition project (1-5 scale)
- Average of the responses for each collaborator in the last six months (1-5 scale for strongly disagree to strongly agree). These measures are adapted from Nyaga et al. (2010).
  - We expect this relationship to continue for a long time.
  - The firm is satisfied with:

    - coordination of activities
    - participation in decision making
    - level of commitment
    - level of information sharing
    - management of activities
  - My firm is satisfied with the collaborative relationship in terms of profitability and sales growth.

- (1-5 strongly disagree to strongly agree) I would recommend this collaborator to other firms looking for business collaborations.

- Number of times the firm has referred the collaborator

D.3 Intermediate Outcomes

D.3.1 Additional Business Outcomes

Business Practice Index

Following McKenzie (2017), this measure consists in the proportion of adopted practices out of a list of 22 which range from marketing to record-keeping, from buying and stock control to financial planning used by the firm. This measure is restricted only to firms that are surviving at the time of the survey. The listed business practices are the following:

- Visited at least one of its competitor's businesses to see what prices its competitors are charging

- Visited at least one of its competitor's businesses to see what products its competitors have available for sale

- Asked existing customers whether there are any other products the customers would like the business to sell or produce
- Talked with at least one former customer to find out why former customers have stopped buying from this business
- Asked a supplier about which products are selling well in this business' industry
- Attracted customers with a special offer
- Advertised in any form (in the last 6 months)
- Attempted to negotiate with a supplier for a lower price on raw material
- Compared the prices or quality offered by alternate suppliers or sources of raw materials to the business' current suppliers or sources of raw material
- The business does not run out of stock monthly or more (coded as one if the business has no stock)
- Keeps written business records
- Records every purchase and sale made by the business
- Able to use records to see how much cash the business has on hand at any point in time
- Uses records regularly to know whether sales of a particular product are increasing or decreasing from one month to another
- Works out the cost to the business of each main product it sells
- Knows which goods you make the most profit per item selling

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

Business Financing Index

We construct this index using the measures listed below following the definition described in McKenzie and Puerto (2020):

- Received at least 1 loan from any source in the past six months
- Received a loan from a bank, microfinance organization, or NGO in the past six months
- Total amount of loans received from all sources in the last six months. 0 if no loans received
- Received a new investment in the form of equity in the past six months
- Has a business bank account that is separate from personal bank account

Capital and Labor Index

We construct this index using the measures listed below following the definition described in McKenzie (2017):

- Total employment: the number of paid workers in the firm, including the owner. Unpaid workers are not included. Coded as zero if the business does not exist
- Value of inventories: current value reported of inventories and raw materials, top-coded at the 99th percentile
- Capital stock: current value of inventories plus the sum of the value of capital purchases made in midline and endline, truncated at the 99th percentile
- Made a large capital purchase: reports making a capital purchase of more than 2000 GHS in the past six months

D.3.2 Business Networking

We construct this index using the measures listed below:

- 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 99th percentile.
- Number of Referrals Received: Number of referrals received in the last six months, truncated at the 99th percentile. Coded as 0 if don't know.
- Member of a Business Association: The firm reports being a member of a business association

D.3.3 Outcomes Related to Access to Legal Support

Business Formalization Index

- Registration documents

- % wage/salaried employees with written contracts
- % employees above the minimum wage
- % suppliers and clients with written contracts
- Ever registering a trademark/patent/copyright

Trust in Contracts

We measure trust in contracts through a trust game modified from a similar survey question used in Cai and Szeidl (2017). Specifically, respondents will be asked the following question:

"Suppose that you are given GHS 10,000. Out of this, you can choose to give as much as you want for a business project which is controlled by another business owner. This project is very successful and triples the money you give. All the proceeds go to the other business owner. The business owner ***[says/agrees in writing that]*** *he will give you 50% of the proceeds. How much (between 0 and GHS 10,000) do you give to this business owner?"* To measure trust in contracts, we randomly vary whether the agreement is verbal or written with equal probability during the survey.

Trust in the Legal System

Respondent answers "agree" or "strongly agree" to the statement "You have high confidence in the legal system".

D.3.4 Ambition and Attitudes

Business Ambitions Index

An average of standardized z-scores for

- number of workers in five years (0 if believe no longer will own business)
- truncated number of workers in five years

- highest monthly sales in five years (0 if believe no longer will own business)
- truncated monthly sales in five years
- expect an increase in sales outside of current market in five years
- expect to export in five years

Get Ahead Attitudes

This index is adapted from McKenzie and Puerto (2020) and is constructed from 11 questions designed to measure positive and optimistic business attitudes. These are scored 1 through 5, where 1=strongly disagree, 5 = strongly agree. Questions will be coded so that higher scores indicate better entrepreneurial attitudes. It include:

- "Even when my business is going well, I keep my eyes open in case I find a way to improve it"
- "When I face a difficult problem, I can usually find some solution"
- "Sometimes I agree to something but then I realize I can't provide it in full or on time, so the customer just has to wait" (negatively coded)
- "I will not try something new unless I am 100% certain it will succeed" (negatively coded)
- "Sometimes to make money you have to risk losing some"
- "I don't worry about where my business will be in the future – I just plan week to week based on what comes up" (negatively coded)
- "If I want to do something, I just do it – I don't need to think about it a lot or discuss with others"

- "I can usually get people to see my point of view, even if they may not understand at first"
- "I am always talking to people and trying to meet new people – you never know when someone will be able to help you later"
- "My business provides about the same as others/is doing about the same as others, so there's no need to make it better" (negatively coded)
- "Even if I am not sure I will succeed in an endeavor, I like to try anyway"

Entrepreneurial Self-Efficacy

We construct an index out of the following 10 business activities that the owner rates themselves as "very confident" in their ability to do. This measure is based on the definition used in McKenzie (2017).

- 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 a 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

Female Empowerment Index

We construct this index using the measures listed below following the definition described in McKenzie and Puerto (2020):

- Compelled to spend money on husband or family (coded 1 if they answer no)
- Not the only person with access to their firms' money (coded as 1 if only they have access)
- Has some money which they have sole control over and can spend how they like
- Do not need anyone's permission:
  - to visit a friend
  - to travel to sell a business asset
  - to travel to a new location to work
  - to stay overnight in a different town
  - to work later than usual hours
  - to take out a loan
  - to spend money on an investment for their business.

D.3.5 Collaboration Trust Index

This index is constructed using the following measures from Seppänen et al. (2007) and Panayides and Venus Lun (2009). Each measure is coded from 1 to 5 for strongly disagree to strongly agree.

- We trust that our collaborator will keep the promises it makes to our firm
- We believe the information that this collaborator provides us

- We trust this collaborator keeps our best interests in mind
- We find it necessary to be cautious with this collaborator (Negatively coded)

D.3.6 Collaboration Formalization Index

We construct index using the measures described below. These measures are adapted from the scale items designed to measure level of formalization in business collaborations in Daugherty et al. (2006).

- Communication between our company and our collaborator takes place frequently.
- The basic terms of our relationship have been explicitly verbalized and discussed.
- We share proprietary information with each other.
- We include each other in formal business planning meetings.
- We have a written agreement.

D.4 Heterogeneous Effects

Business Networking Index

- 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 99th percentile
- Number of Referrals Received: Number of referrals received in the last six months, truncated at the 99th percentile. Coded as 0 if don't know
- Member of a Business Association: The firm reports being a member of a business association

Legal Knowledge Score Index

Index of 12 questions designed to measure the level of legal knowledge

- Which is true about an employer?
  - An employer must pay taxes and pension contributions for employees
  - An employer does not pay for overtime
  - An employer is different from the business owner
  - Don't know
  - Refused to answer
- What statutory filings are required for limited liability businesses?
  - Annual Returns
  - Directors minutes
  - Business plan
  - Don't know
  - Refused to answer
- True or False: An employer is required to register with the Social Security and National Insurance Trust (SSNIT)
  - True
  - False
  - Don't know
  - Refused to answer
- True or False: My business requires a licence to operate.

  - True
  - False
  - Don't know
  - Refused to answer

- It is safe to secure your loans with what?
  - Personal assets
  - Business assets
  - Family land
  - Don't know
  - Refused to answer

- Who is exempt from paying income taxes?
  - Firms operating in the informal sector (e.g. small-scale self-employed dressmakers / tailors, traders)
  - Individuals with total income below GHS 261 per month
  - Unregistered individuals/firms
  - No one
  - Don't know
  - Refused to answer

- True or False: If you are a young entrepreneur under the age 35 and is a sole proprietor, you can receive an income tax exemption on business income for the first five years: True or false?

  - True
  - False
  - Don't know
  - Refused to answer

- What's the main legal difference between a limited liability company and a sole proprietorship?
  - Limited liability companies must have more than one owner.
  - Sole proprietorship can not fail.
  - In a limited liability company, there is a distinction between business and personal assets so the owner is not personally responsible for business debts.
  - In a sole proprietorship, the owner does not need to pay income taxes.
  - Don't know
  - Refused to answer

- If someone lends you money for your business, that person becomes co-owner of your business. True or false?
  - True
  - False
  - Don't know
  - Refused to answer

- Which of the following protects the names and identifying marks of products and companies?

  - Patents
  - Trademarks
  - Trade secrets
  - Copyrights
  - Don't know
  - Refused to answer

- After registering a business name with the Registrar's General Department, how frequently do you have to renew the registration?
  - Every 6 months
  - Every year
  - Once every 5 years
  - Never
  - Don't know
  - Refused to answer

- True or False: If a (non-casual) worker is employed for longer than six months, a written contract must be provided.
  - True
  - False
  - Don't know
  - Refused to answer