

Why Don't Small Firms Merge?

Experimental Evidence on Information Barriers

Pre-Analysis Plan

Morgan Hardy, Seongyoon Kim, Andreas Menzel, Marc Witte

March 28, 2021

1 Introduction

Compared to firms in developed countries, those in developing countries are smaller, with many of them operating as single-person businesses (Hsieh and Olken, 2014). Some of these firm owners prefer to be an entrepreneur over having fixed wage employment, while others prefer wage employment and run their businesses as a necessity (De Mel, McKenzie and Woodruff, 2019, Karaivanov and Yindok, 2019, Schoar, 2010). One way firms in these contexts could grow is when the "innate entrepreneurs" hire and absorb the firms of the many "necessity entrepreneurs".

This research project aims to experimentally investigate information barriers to small firm consolidation. We exploit rich panel data on the universe of garment making firm owners in Hohoe, Ghana, including owner and owner household characteristics, firm characteristics and outcomes, and reported within-industry peer network interactions. We elicit firm owners' willingness to pay for information about within-industry peer willingness to hire and be hired by the owner using Becker-DeGroot-Marschak method (Becker, DeGroot and Marschak, 1964) and exploit the conditionally random variation to

study the impacts of such information on firm-level growth and profitability as well as dyadic-level mergers.

This document describes the Pre-Analysis Plan for the experiment. This draft was completed and registered after the data collection, but prior to the analysis of any follow-up data. Therefore, it provides a useful reference in evaluating the final results of the study. Section 2 describes the project data. Section 3 describes the experiment design. Section 4, 5, and 6 present our analysis plan for willingness-to-pay for information, firm-level outcomes, and dyadic-level outcomes, respectively.

2 Data

This research project exploits data on a comprehensive sample of tailors in Hohoe, Volta Region, Ghana. We conducted a full census of all firms in the garment sector based in Hohoe in August/September 2019.¹ As of September 2019, our census identified 569 garment makers in Hohoe. Immediately following the census, we collected a baseline survey with each firm owner in October/November of 2019 that included a full network mapping activity. Our willingness-to-pay exercise and information randomization was rolled out in January 2020. Finally, we conducted an endline retrospective phone survey between August 2020 - January 2021 (with questions covering February to July of 2020).

2.1 Census

The census data include basic information, such as firm owner name, gender, firm location, contact details, the number of garment makers they know in the district, the number of garment makers they know outside the district, number of workers, number of years in business, and a photo, which was later used for respondents to identify other garment

¹We define 'garment maker' as someone who was operating a garment making business for an income at the time of the census. The census began with a followup on a similar census sample identified in 2014 and was completed by snowball sampling and block canvassing.

makers.

2.2 Baseline

Our baseline survey data collection consists of two components: A) a respondent-level survey and B) a network survey. The respondent-level survey captures a plethora of characteristics of the firm owner (ex: age, ethnicity, schooling, cognitive skills, management score, reasons for self-employment), household (ex: size, employment, finances) and business (ex: asset value, revenue/profits last month, self-reported barriers).

To begin the network survey, we ask respondents to list all garment makers in Hohoe with whom they have interacted during the past year and match each contact mentioned to the specific sample member using our census book. For each of the the identified contacts, we ask respondents questions that characterize the relationship including how often they interact, how long they have known each other, and whether they have outsourced to each other.

After basic relationship information, respondents participated in a hypothetical (not cash incentivized) trust game for every identified contact. The trust game is often used in the literature to measure interpersonal levels of trust. We ask respondents what they would do if they receive 10 Ghana Cedis (1.80 USD at the time of the experiment) and the contact person receives 0. They could choose to give 0 GHC to the person, which leaves the respondent with 10 GHC, and the contact person with 0 GHC, or to give 10 GHC to the person, which would be tripled to the person, leaving the contact person with 30 GHC. Of the 30 GHC, the contact person can return some of that money to the respondent. In addition to each identified contact, this hypothetical game was also repeated for “the average garment maker you do *not* know in Hohoe”. This question attempts to sort pairs into high levels of trust and low levels of trust.

Importantly, we also ask respondents about their willingness to hire every identified contact and, if willing, their reservation pay (the highest amount they’d be willing to pay for the labor) as well as their willingness to work for every identified contact and, if

willing, their reservation wage (the lowest amount for which they'd be willing to supply their labor). These questions were asked about each i) outsourcing, ii) piece-rate contracts, and iii) fixed weekly wage contracts.² Each respondent was randomly assigned to one of two groups, where the treatment group was asked to answer these questions under the hypothetical scenario that the employer of the contract has “consistently many garment orders for three months”. In addition to each identified contact, these questions were also repeated for “the average garment maker you do *not* know in Hohoe”. These questions are used to populate the lists of names offered as part of the information intervention described below.

2.3 Willingness to Pay Exercise and Information Roll-out

We conduct a two-module survey in January 2019. In the first module, we ask respondents time-varying questions regarding the household and their business, similar to the baseline. In the second module, we measure respondents' willingness to pay for information about other garment makers in their district who are willing to work for them and/or are willing to hire them.

Using responses from willingness to work and willingness to hire questions, we randomly generate four different lists that consist of contact details of seven firm owners. The information that each list provided differed in two ways: willingness to hire versus willingness to work, and whether the seven firm owners were ordered by past sales or how their past sales negatively correlated with the respondent's. The lists are shown below:

- “List A: a list of 7 other tailors that have indicated a willingness to *work for you* on a fixed weekly wage in decreasing order of their past sales”
- “List B: a list of 7 other tailors that have indicated a willingness to *work for you* on a fixed weekly wage in decreasing order of negative correlation of expected sales” (to

²These questions can be found in Appendix A

smooth income streams from the combined businesses)

- “List C: a list of 7 other tailors that have indicated a willingness to *hire you* on a fixed weekly wage in decreasing order of their sales”
- “List D: a list of 7 other tailors that have indicated a willingness to *hire you* on a fixed weekly wage in decreasing order of negative correlation of expected sales”

Following the Becker-DeGroot-Marschak (BDM) method, each respondent receives 5 GHC (approximately 91 US cents at the time of the first wave) that can be used as offers for the lists of other garment makers that are willing to work for or willing to hire the respondent. Respondents can then bid for each list (in integer amounts), with a maximum possible bid for each list of 3 GHC. Then, the price for the list is randomly drawn and takes integer values between 0 and 4 GHC. If the bid is greater than or equal to the price, then the list is successfully purchased, and any amount not spent is returned to the respondent in cash. Each respondent was also placed into four different treatment groups, where the information on the list varied. More details on the content of these lists and the treatment groups are provided in the next subsection.

2.4 Endline

Beginning in August 2020 going through January 2021, we collect endline data via phone surveys given the COVID-19 pandemic. For every month between February and July 2020, we collect retrospective data on key business outcomes. These variables include sales, wage expenses, profits, and hours worked in average week. In addition, for the two months of February and July, we also capture other income-generating activities, household consumption, and a full labor roster with anyone who contributed some work in some form. We plan to use name-matching techniques to see if some garment makers in our sample have worked for other garment makers. Finally, we also ask questions regarding the challenges that the respondents’ business faced during the COVID-19 pandemic.

3 Experimental Design

In the second wave of data collection, we ask respondents to participate in a Willingness-to-Pay exercise where respondents bid for lists of information on other garment makers interested in firm consolidation, as stated above. Each respondent was also randomly assigned into one of four different treatment groups. Each treatment group was given the choice of the same four lists, but the information these lists provided differed in the following way:³

1. Treatment Group 1: “...willing to work for you/hire you on a PIECE RATE, IF YOU HAD RELIABLY MANY ORDERS ...”
2. Treatment Group 2: “...willing to work for you/hire you on a PIECE RATE ...”
3. Treatment Group 3: “...willing to work for you/hire you on a FIXED WEEKLY WAGE, IF YOU HAD RELIABLY MANY ORDERS ...”
4. Treatment Group 4: “...willing to work for you/hire you on a FIXED WEEKLY WAGE ...”

Note that the lists purchased were generated and distributed in the weeks immediately after these willingness-to-pay exercise. This is because the willingness-to-pay exercise was prefaced by an additional consent module. Opting in implies that respondents are willing to participate in the elicitation of willingness to pay and *also agree to have their basic contact information potentially shared with other firm owners*. Only those who opted in were included on lists. The majority of firm owners reached opted in. All respondents who opted into this part of the study had more than seven firms that could populate each list. From these groups, seven firms were randomly selected for each list for each respondent, and then ordered according to the information given in the list.

³The treatment assignments for Wave 2 are further illustrated in the appendix.

4 Willingness-to-pay Analysis Plan

In a first step, we analyze whether the *type* of information about other garment makers that we offer respondents affects their willingness-to-pay for that information. Here, we make use of the random variation in the type of information on the lists specified in the previous section.

$$WTP_i = \beta_0 + \beta_1 FwD_i + \beta_2 FwND_i + \beta_3 PrD_i + \epsilon_i \quad (1)$$

The variables in equation (1) are defined as follows:

- WTP_i is respondent i 's Willingness-to-Pay for a given type of information: list A, B, C, D, or pooled.
- FwD , $FwND$, PrD are treatment indicators. They refer to the Fixed Wage-Demand, Fixed Wage-No Demand, and Piece Rate-Demand treatment groups. The fourth treatment group (Piece Rate-No Demand) is the omitted comparison group.

4.1 Heterogeneity

The Willingness-to-Pay analysis specified in equation (1) might be affected by firms' base-line characteristics and preferences. To identify such heterogeneity, we pre-specify the following sources of potential heterogeneity in Willingness-to-Pay.

- Do firm owners prefer information that reveals volume of sales over negative correlation of sales?
 - Respondents may have preferences on whether they want the lists to reveal information on whose sales are the greatest or whose sales correlate negatively with theirs. This may reveal preferences on whether respondents want to safeguard themselves against demand shocks and smooth income.

- Do firm owners prefer information on firm owners that are willing to work for or willing to hire them?
- Does treatment effect vary between firm or firm owner characteristics?
 - Gender
 - Baseline sales (or expected sales)
 - Sales volatility (or expected sales volatility)

5 Firm-Level Outcomes Analysis Plan

In a second step, we analyse whether the purchase of a specific list affects downstream firm outcomes. Firms were offered to purchase the various lists based on the Willingness-to-Pay exercise described in section 3. We analyze the effect of receiving a list with information about other garment makers on own firm outcomes with the following specification:

$$y_i = \beta_0 + \beta_1 W_i + \beta_2 n(W)_i + \beta_3 H_i + \beta_4 n(H)_i + \gamma_1 T_i \times \overline{WTP}_i + \gamma_2 N(W)_i + \gamma_3 N(H)_i + u_i \quad (2)$$

The variables in equation (2) are defined as follows:

- y_i is firm i 's outcome of interest
- W_i indicates whether respondent i purchased a list of tailors who have indicated a **willingness to work** for i .
- H_i indicates whether respondent i purchased a list of tailors who have indicated a **willingness to hire** i .
- $n(W)_i$ is the **number of tailors willing to work** for i , who respondent i identified by purchasing lists.

- $n(H)_i$ is the **number of tailors willing to hire i** , who respondent i identified by purchasing lists.
- Controls:
 - $N(W)_i$ is the number of willing *workers* who were eligible to be on i 's list (i.e. all workers who potentially could have been on the list)
 - $N(H)_i$ is the number of willing *employers* who were eligible to be on i 's list (i.e. all employers who potentially could have been on the list)
 - \overline{WTP}_i refers to the unique combination of Willingness-to-Pay for each of the lists A, B, C, D for firm-respondent i
 - T_i refers to the treatment assignment for respondent i (fixed wage/piece rate, demand/no demand).

While equation (2) specifies a linear parametric form, we will also run alternative versions of the model in which we include $N(W)_i$, $N(H)_i$, $n(W)_i$, and $n(H)_i$ non-parametrically. Furthermore, we are also interested in testing additional interaction effects between the main explanatory variables of interest in equation (2), such as the interaction between W_i and H_i .

5.1 Heterogeneity Analysis

Following the main firm-level analysis, we will conduct a heterogeneity analysis, focusing on the following main dimensions of heterogeneity:

1. Do firms need assurance of “reliably many orders” before hiring or working for other respondents?
 - Treatment groups 1 and 3 are given lists with the phrase “if you had reliably many orders” and treatment groups 2 and 4 are not. We plan to investigate whether this additional information significantly affects the treatment effect.

2. Does it matter whether the information is based on piece-rate contracts or fixed-wage contracts?

- Treatment groups 1 and 2 are given lists of people willing to work for or hire the respondent based on piece-rate contracts, while treatment groups 3 and 4 are given lists based on fixed-wage contracts. It is possible that the treatment effect varies depending on what kind of contracts the information treatment is based on.

3. Does the treatment effect vary based on firm or firm owner characteristics?

- Gender
- Baseline sales (or expected sales)
- Sales volatility (or expected sales volatility)
- Interpersonal trust

In line with the heterogeneity analysis, we will also conduct the following additional pieces of analysis:

- Interact heterogeneity variable of interest with all $W, n(W), H, n(H)$.
- Pool the WTP for list A and list B, lists C and D, and interact these with the heterogeneity variable of interest.

5.2 Outcome Variables

Here, we briefly present a an overview of our firm-level outcome variables of interest, classified into primary, intermediate and secondary outcomes.

- Primary outcome variables
 - Ownership (closure, co-ownership)
 - Firm performance (profits, revenues)

- Intermediate outcome variables
 - Firm size (number of workers, wage expenses)
 - Firm boundaries (outsourcing revenue and expenses, capital rental revenue and expenses)
- Secondary outcome variables
 - Income-related (number of other income-generating activities and total income, financial liquidity)
 - Welfare-related (consumption, household assets, COVID-19 Impact)
 - Household-related (income from household members in public, private sector, farming, or self-employed)

5.3 Multiple hypothesis testing

We will correct for multiple hypothesis testing in two ways. First, whenever possible we use summary measures of relevant concepts as main outcomes whenever possible. Second, we will perform robustness checks that adjust standard errors (for example, q-values) to correct for multiple hypothesis testing within each family of non-main outcomes. However, since our key outcomes of interest are pre-specified and specific, we do not think these adjustments should be a first order concern.

5.4 Attrition

We will track attrition rates at followup. We will use Lee bounds if attrition imbalance is found across treatment groups.

6 Dyadic-Level Outcomes Analysis Plan

In a third and final step, we will analyze the effect of our information treatment on dyadic outcomes, i.e. outcomes between two tailors in our sample. We analyze the effect of receiving a list with information about other garment makers on dyadic outcomes with the following specification:

$$y_{ij} = \beta_0 + \beta_1 W_{ij} + \beta_2 H_{ij} + \gamma_1 P_{ij} + \gamma_2 P_{ji} + \gamma_3 T_i \times \overline{WTP}_i + \gamma_4 T_j \times \overline{WTP}_j + \epsilon_{ij} \quad (3)$$

The variables in equation (3) are defined as follows:

- y_{ij} refers to the dyadic outcome variable of interest between firms i and j .
- W_{ij} refers to individual i purchasing a list of potential employees with j on it.
- H_{ij} refers to individual j purchasing a list of potential employers with i on it.
- Controls:
 - P_j refers to the probability that j is selected to be one of the seven firm owners on the list that individual i may purchase.
 - P_i refers to the probability that i is selected to be one of the seven firm owners on the list that individual j may purchase .
 - T_i is a treatment indicator for individual i
 - \overline{WTP}_i refers to the unique combination of Willingness-to-Pay for each of the lists A, B, C, D for respondent i .

Note: The control variables, namely P_i and P_j will both be analyzed in parametric form with a linear and/or square term, and in non-parametric form with indicator variables of all possible probabilities.

6.1 Heterogeneity Analysis

Following the main specification of the dyadic analysis, we will analyze potential sources of heterogeneity. We will do this by interacting the heterogeneity variable of interest with the main explanatory variables W_{ij} and H_{ij} . We will focus on the following main dimensions of heterogeneity:

1. Type of contract that i and j worked together on (outsource, piece-rate, fixed wage)
2. Whether i knows j , j knows i , or they both know each other
3. Whether i 's list type matches j 's list type (piece rate list and piece rate list, or fixed wage list and fixed wage list)
4. Whether j was on i 's list as well as i being on j 's list
5. Homophilic characteristics
 - Network distance (as measured by shortest path between i and j)
 - Geographical distance
 - Co-ethnicity
 - Same gender
 - Years they have known each other

6.2 Outcome Variables

Here, we briefly present a an overview of our dyadic outcome variables of interest, classified into primary and secondary outcomes.

- Primary outcomes of interest
 - y_{ij} for those tailors i with lists of tailors willing to work for them
 - * Did garment maker j work for i ?

- * Did garment maker j outsource or rent capital to garment maker i ? How much?
- y_{ij} for those tailors i with lists of tailors willing to hire them
 - * Did another garment maker j hire tailor i ?
 - * Did garment maker j receive outsource work or did j rent capital from garment maker i ? How much?

References

- Becker, Gordon M, Morris H DeGroot, and Jacob Marschak.** 1964. "Measuring utility by a single-response sequential method." *Behavioral science*, 9(3): 226–232.
- De Mel, Suresh, David McKenzie, and Christopher Woodruff.** 2019. "Labor drops: Experimental evidence on the return to additional labor in microenterprises." *American Economic Journal: Applied Economics*, 11(1): 202–235.
- Hsieh, Chang-Tai, and Benjamin A Olken.** 2014. "The missing" missing middle"." *Journal of Economic Perspectives*, 28(3): 89–108.
- Karaivanov, Alexander, and Tenzin Yindok.** 2019. "Involuntary Entrepreneurship: Evidence from Thai Urban Data."
- Schoar, Antoinette.** 2010. "The Divide between Subsistence and Transformational Entrepreneurship." *Innovation Policy and the Economy, Volume 10*, 57–81. University of Chicago Press.

Supplementary Appendix

A Firm Merger Questions - Baseline

Phrases in italics are for the treatment group. The control group was asked these questions without the italicized phrases.

- *If you know that you would have consistently many garment orders every week for three months*, is there a wage at which you would outsource all or some of an order to this person? What is the highest amount you would offer this person per child's shirt?
- *If you know that you will have consistently many garment orders every week for three months*, and if you need an additional piece-rate worker, is there a wage at which you would be open to having the person work for a piece rate in your shop? What is the highest piece-rate amount you would offer this person per child's shirt?
- *If you know that you will consistently have many garment orders every week for three months*, and if you need an additional fixed weekly wage rate worker, would you be open to having the person work for a fixed weekly wage rate in your shop? What is the highest weekly payment you would offer this person to work for you?
- *If you know that this person will have consistently many garment orders for three months*, is there a wage at which you would want to be outsourced all or some of an order from this person? What is the lowest amount you would take per child's shirt to work for this person?
- *If you know that this person will have consistently many garment orders for three months*, is there a wage at which you would be open to work for this person for a piece rate at their shop? What is the lowest amount you would take per child's shirt to work for this person?

- *If you know that this person will have consistently many garment orders for three months, is there a fixed weekly wage at which you would be open to work for this person for a fixed weekly wage rate in their shop? What is the lowest amount per week you would need to take to work for this person?*

B Experimental design

Below, Figure A1 highlights the four treatment groups respondents are pre-assigned to in the second wave when we elicit willingness-to-pay.

Figure A1: Wave 2 Treatment Groups

