Competition, Efficiency and Trust: Evidence from Rwanda's Coffee Wet Mills

Rocco Macchiavello — Warwick University
Ameet Morjaria — Harvard University

August 2014*

PRELIMINARY & INCOMPLETE
NOT FOR FURTHER CIRCULATION

Abstract

We examine the role of competition on productivity and management practices in the context of coffee wet mills in Rwanda. The sector is characterized by (i) simple technology which allows for a transparent measurement of efficiency, (ii) exogenous determinants of entry and (iii) a common set of appropriate management practices. We document substantial cross-sectional dispersion in the efficiency of wet mills. We build upon an engineering model for the optimal placement of coffee wet mills and a comprehensive GIS database to instrument for the level of competition in coffee procurement. Conditional on cost drivers at and around the wet mill, the competition faced by the stations in sourcing coffee cherries strongly correlates with the score derived from the engineering model around the mill's catchment area. While competition leads to moderate increase in coffee cherry prices, imperfect enforcement in the market implies that competition i) reduces capacity utilization and efficiency, ii) increases unit labor and finance costs, iii) reduces trust between the station and the farmers, iv) does not lead to higher fertilizers use and yields among farmers. The evidence suggests scope for inefficient entry in the industry and rationalizes commonly observed policies in agricultural markets, such as zoning regulations.

Keywords: competition, productivity dispersion, trust, pecuniary externalities.
JEL Codes: L14, L22, O12, Q13.

*We are grateful to several staff members at the Rwanda National Agricultural Exporting Board (NAEB) for their support. In particular, the former Director General Alex Kanyankole, Eric Rukwaya, Laetitia Umotoniwase, Vincent Uwimana, Badrou Bazambanza, Celestin Gatarayiha and Faustin Nizeymana. We also thank the IGC team, especially Jonathan Argent, Sophia Kamaruddin and Richard Newfarmer for their support. Morjaria thanks NAEB and MinAgri for hosting his stay in Rwanda. Financial support from IGC Rwanda Country Program is gratefully acknowledged. We thank seminar and workshop participants in Ascea Summer School, DIME-CEGA Kigali Workshop, Goteborg, IIES, ISNIE 2014 (Duke), Oxford and Warwick for helpful comments and suggestions. The views expressed here are those of the authors alone and do not necessarily reflect those of DFID, NAEB, Ministry of Agriculture or agencies that shared data. Corresponding author: Morjaria, Harvard Academy for International and Area Studies, email: ameet.morjaria@gmail.com

1 Introduction

A substantial empirical literature documents large, persistent productivity differences (PPDs) across firms, even within narrowly defined sectors (see, e.g., Syverson 2011). To a certain extent, PPDs must originate from imperfect competition that allows inefficient firms to survive in the market place (Syverson 2004). This explanation resonates with the observation that productivity differences appear to be even more pronounced in lower income countries, where a variety of institutional and regulatory frictions, e.g., regulatory barriers to entry and financial constraints, prevent more efficient firms to expand (Hsieh and Klenow, 2009). Indeed low competition is associated with poor management practices (Bloom and Van Reenen, 2007) and new evidence from India appears to show this is also associated with lower productivity (Bloom *et al* 2013).

On the other hand, when markets are highly imperfect, as might often be the case in developing countries, competition might have perverse effects on productivity. A tendency towards excessive entry naturally arises in a variety of contexts (see, e.g., Mankiw and Whinston (1986)), for example in models of search (see, e.g., Diamond (1982), Hosios (1990)) or when implicit arrangements are necessary to sustain trade (e.g., Kranton (1996) and Gosh and Ray (1996), Fafchamps (2004)). These mechanisms are likely to be particularly relevant in the case of agricultural value chains, which tend to be plagued by numerous market imperfections (Binswanger and Rosenzweig (1986)). As a result of these market imperfections, governments have often regulated entry and competition in agricultural markets, e.g., establishing monopsony rights and zoning regulations (see, e.g., World Bank (2008), Mullainathan and Sukhtankar (2014)).

Providing empirical evidence on the effect of competition on efficiency is, however, challenging since (i) precise measurement of both productivity and competition are rarely available, and (ii) entry of competitors is likely to be a result as much as a cause of a firm's poor management and performance. This papers studies the relationship between competition and efficiency in the context of coffee wet mills in Rwanda. We focus on competition in the input markets, i.e., how mills compete to source coffee cherries during the harvest season. The sector offers a number of vantage points to make empirical progress. First, the sector is characterized by a simple and uniform technology that allows precise measurement of (drivers of) productivity and unit costs, as well as appropriate managerial practices. Second, we build upon a detailed engineering model for the optimal placement of wet mills in order to construct an instrument for the competitive pressures faced by individual washing stations. The engineering model relies on detailed spatial data vectors and predicts very well the actual placement of

mills in the data. The model, therefore, allows us to construct a measure of suitability for the placement of a washing station at a very localized level (one square kilometer). Controlling for the placement score as well as a broad vector of factors affecting coffee cultivation within the mills catchment areas, we use the predicted placement in areas *outside* the catchment area to instrument for the competition actually experienced by mills. To complement the information available from administrative data collected by the coffee board on capacity, production and prices of the mills, we conducted an original survey of essentially all wet mills (178) that were operating in Rwanda during the 2012 harvest season.

We begin our empirical analysis by document substantial cross-sectional dispersion in the efficiency of wet mills. On average, with variation depending on geographic and soil characteristics, it takes approximately 6 kilograms of coffee cherries (inputs) to produce 1 kilogram of parchment coffee (output). We follow standard practices in the industry and measure efficiency as the amount spent to process one kilogram of parchment coffee. Once purged by substantial cross-sectional variation induced by soil and geographic characteristics around the mill, we find the ratio of the 90^{th} and 10^{th} percentile in the distribution of unit costs to be approximately 1.5. Since the costs of cherries account for approximately 60% of the unit costs, the relatively low ratio conceals substantially higher heterogeneity in the part of costs directly under control of managerial practices - capital and labour - for which the 90^{th} to 10^{th} percentile ratio is 2.5.

In line with theoretical arguments emphasizing pecuniary externalities caused by competition, we find that competition, as measured by the number of competing mills sourcing coffee in areas very close to the mill, positively correlates with unit costs. Conditional on a host of soil and geographical controls around the mill we find that competition is associated with mildly higher coffee cherry prices received by the farmers for their produce. This effect is, obviously, in line with the predictions of perfect markets in which competition translates into higher costs by affecting prices. But competition is also associated with significantly higher unit costs, other than those directly related to the purchase of coffee cherries. In particular, competition correlates with higher labour and capital unit costs. A number of additional correlation suggest that competition affects these costs by inducing lower (and more volatile) capacity utilization, fewer interlinked transactions with farmers (e.g., purchases on credit, second payments, loans for fertilizers) and, crucially, lower level of trust between the farmers and the station.

We then utilize knowledge of the engineering model to predict the optimal placement of coffee wet mills. We construct a comprehensive GIS database to estimate the engineering model. We find that the model predicts very well placement of mills. We then use the estimated engineering model to instrument for the level of competition in coffee procurement as follows. Conditional on cost drivers at and around the wet mill, we instrument for competition using the predicted score from the engineering model in areas around the mill's catchment area. The first stage between the instrument and the competition level reported by the firm is strong and highly significant. In the second stage, the IV results confirm the results in the OLS analysis. First, competition only mildly increases input prices. Second, due to imperfect enforcement in the input market, competition also (i) reduces trust between the station and the farmers, (ii) reduces capacity utilization and (iii) increases both labour and capital unit costs. In equilibrium, competition makes it harder for stations to operate and, as a result, the total amount of cherry purchased by stations and prices paid to farmers are not significantly higher than what would have prevailed with lower competitive pressures. Using both a representative survey of farmers around the surveyed mills as well as a comprehensive census of coffee farmers carried out in 2010, we find no evidence that competition leads to higher yields of fertilizers use among farmers.

The rest of the paper is organized as follows. Section 2 discusses related literature. Section 3 provides background information on the industry and describes the data utilized in the analysis. Section 4 presents our measure of efficiency and unit costs and documents dispersion in efficiency levels across firms. Section 5 introduces a simple theoretical framework which guides the empirical exercise. Section 6 presents the main results. The section first presents the OLS specifications results between efficiency and competition. Then, the entry model and the construction of the instrumental variable for competition are described. Finally, the IV results are presented. Section 7 empirically explores the mechanisms behind the main results as well as the effects of competition on farmers. Section 8 offers some concluding remarks.

2 Related Literature

[TO BE ADDED]

3 Background

Sector: Coffee in Rwanda

The coffee sector has been an important contributor to the Rwandese economy for several decades. Coffee, which until the genocide was Rwanda's main export, became widespread in the late 1930s following five waves of mandatory coffee-tree planting imposed by the Belgian colonial administration in order to increase revenue collection from its Rwanda-Urundi colonies in the context of the Great Depression (see, e.g., Blouin (2014)). At the time of independence, coffee represented 55% of Rwandese exports against minerals (37%), pyrethrum (3%) and tea (2%).The Rwanda's exports sector almost collapsed between 1986 and 1995. The decline started in the mid '80s and was accelerated by the coffee crisis which followed the demise of the International Coffee Agreement in 1989 (and the subsequent collapse of coffee prices in the global market). This decline was amplified by the un-sustainability of inward oriented policies and eventual economic instability following consecutive devaluations in the early 1990s and the political instability leading to the 1994 genocide. Following the end of the genocide, the sector has steadily recovered. Today, there are around 400,000 farmers, mostly small holders, growing coffee in Rwanda and coffee accounts for almost 40% of Rwandese exports. The number of coffee washing stations, which perform the first stage processing as further described below, has increased from only one station active in 2001 to more than 200 stations active during the last harvest in 2014 (see Figure 1). Today, it is estimated that the coffee sector as a whole accounts for about 12% of the entire gross domestic product of the country.

Data Description

Wet Mill Census 2012: The empirical analysis uses survey data which was collected through a census of wet mills. The survey was designed by the authors in collaboration with the National Agricultural Exporting Board and was implemented between May and July 2012. The survey aimed to cover all operating coffee wet mills in the 2012 season, four survey team's each including six team members were trained by one of the authors to administer the survey and cover the whole of Rwanda over the season. Upon previous appointment, face-to-face interviews were conducted, the average survey completion time was 5 hours per wet mill, on average a team was able to survey two wet mills a day. Modules included surveying (i) the manager (and owner if necessary) and covered general questions about the wet mill (separate modules for private or cooperative wet mills, wet mill history, assets owned by the mill, donor assistance,

operational activity, procurement of cherries, relationship with farmers, employment, access to finance, marketing and capital invested) and than a specific module focused on interviewing the manager (socio-economic characteristics, canonical World Value Survey questions, employment history, earnings from station and income profile, incentive payments, raven and numerical test), (ii) the main collector of cherries in the locality for the wet mill (this included conducting raven and numerical tests, questions on the socio-economic characteristics, history of the relationship with the manager and the wet mill as well as their asset ownership), (iii) random selection of five farmers (chosen by NAEB district officer, survey team choose random letter for last name and provided to district officer in advance of reaching the wet mill) who were asked about their socio-economic characteristics, raven and numerical tests were conducted, various questions related to their coffee production (ownership of trees, past and current production and sales), expectation on prices, membership in cooperatives, interlinked transactions, input and canonical trust questions and (iv) four random workers were also interviewed at the wet mill, who were chosen on the spot by the survey team on arrival at the mill, questions covered, raven and numerical tests, canonical World Value Survey questions, socio-economic questions, earnings from wet mill and relationship with the manager. Sample of coffee lots were also taken from each wet mill (mills were advised to provide their best lot) to be physically examined and cupped in NAEB's laboratory in Kigali and lastly a GPS module of the wet mill was conducted (longitude, latitude, elevation).

Administrative and Geographic Information System Data: Several ancillary data sources were obtained from NAEB as well as other agencies (Ministry of Agriculture, Rwanda Natural Resources Authority, Ministry of Infrastructure, Rwanda Meteorological Agency, TechnoServe, National University of Rwanda - Butare). GIS tools allows us to divide Rwanda into a one square kilometer grid (Rwanda's size is circa. 26,000 square kilometer). We are than able to assign each square kilometer, grid box, its geographic characteristics (altitude, slope, historical coffee suitability, different types of roads and rivers and rainfall, we run our analysis both at the grid and wet mill level. We also have access to a detailed coffee census conducted in 2009 that surveys all coffee farmers (circa. 400,000 farmers) to understand the welfare impacts due to entry of wet mills, as the census asks retrospective questions for the period 2007-2009.

4 Unit Costs and Efficiency Dispersion in the Industry

Coffee Wet Mills

The coffee cherry is the fruit of the coffee plant. The cherries are ripe when they change color from green to red, at which point they should be harvested. The harvest season typically lasts for three to four months. The timing of the harvest season varies by country and, within country, by region depending on altitude, soil and rainfall patterns. Most coffee-growing countries have only one major harvest a year. In most countries, the coffee cherries are picked by hand, a labor-intensive and difficult process.

The pulp of the coffee cherry is removed leaving the seed or bean which is then dried to obtain parchment coffee. There are essentially two processing methods to obtain parchment coffee: the dry method and the washed method. The dry method, also known as unwashed or natural, has the entire cherry first cleaned and then placed in the sun to dry on tables. This process is done by farmers. In the washed method, instead, farmers bring coffee cherries immediately after harvest to a washing station, the firms that are the objects of our study. The washed method requires specific equipment and substantial quantities of water. After the cherry skin and some of the pulp are removed with a pressing machine, cherries are then sorted by immersion in water. The bean is then left to ferment, typically for around 30 hours, to remove the remaining skin. The fermentation process has to be carefully monitored to ensure that the coffee doesn't acquire undesirable, sour flavors. When the fermentation is complete, the coffee is thoroughly washed with clean water in tanks or in special washing machines. The beans are then dried in the sun (and sometime with the further help of machines). Coffee dried in the sun is spread out in rows on large tables or on patios where it is frequently turned by hand or raked. Drying coffee on tables improves quality but increases cost and labor significantly. After the drying process the parchment skin is easily removed in the hulling process. Essentially all coffee is hulled before shipment.

The method used to process the cherries has a significant effect on the flavor of coffee once roasted and brewed. The wet method delivers higher consistency and quality which is reflected in prices. Data from customs records and the survey we conducted in Rwanda, for example, reveal that washed coffee is sold at a substantial price premium (around 40%) relative to dry coffee both as parchment in the domestic market and as green coffee at the export stage.

Descriptive: Firm's Size

Despite having seasonal activities tied to the coffee harvest season, wet mills can be

thought of as large firms by developing country standards. For example, Blouin and Macchiavello (2013) study a sample of 200 ca. coffee washing stations in 20 countries and document that the average stations has sales in excess of $1.0 million a year, holds assets also in excess of $1.0 million, and hires more than 100 employees during the harvest season. These figures makes coffee washing stations large, given the typical firms size distribution in developing countries (see Hsieh and Olken (2014)). Relative to the sample in Blouin and Macchiavello (2013), coffee washing stations in Rwanda appear to be somewhat smaller on average, with average employment during harvest season of 67 employees. This still leaves coffee washing stations comfortably in the right tail of the firm size distribution in Rwanda (see, e.g., Söderbom and Kamarudeen (2013)). Still, the average coffee washing station in Rwanda purchases from close to 300 small holders farmers in their surrounding area.

The sector is Rwanda is characterized by three different types of firms: i) privately held mills, ii) large cooperative mills and iii) TechnoServe (cooperative) mills. There are approximately 80 privately held mills, 80 large cooperative mills and 40 TechnoServe mills. As can be seen in Table 1, the three different types of firms differ along a number of dimensions. In general, relative to cooperative and TechnoServe mills, private mills tend to be larger, have lower utilization rates, higher unit costs but more professional managers. These differences, however, are largely driven by the TechnoServe mills which operate a significantly smaller technology and have, therefore, lower unit costs and higher utilization rates.

Figure 2 illustrates the firm size distribution according to installed capacity as well as capacity utilization.[1] The left panel documents relatively large dispersion in capacity, with TechnoServe mills having capacity installed of approx. 100 or 150 tons, the bulk of the mills being medium sized with capacity of 500 tons and very few (private) mills with much larger capacity. The right panel documents huge dispersion in utilization rates during the 2012 harvest season. Besides inactive mills at zero capacity utilization, the median mill has a capacity utilization well below 50%.[2]

Descriptive: Unit Costs

Coffee washing stations are characterized by a relatively simple technology. In particular, it takes approximately 5.5 Kgs of coffee cherries (inputs) to produce 1Kg of

[1] These figures are preliminary.

[2] At the time of the survey, total installed capacity in the sector would have allowed the country to export approx. 70-80% of harvest as fully washed coffee. However, due to low utilization rates, only 30-40% of the coffee was exported as fully washed. Given a 40% price premium, this amounts to a sizable loss in export revenue for the Rwandese economy.

parchment coffee (output).[3] Together with the fact that, typically, the direct costs of purchasing coffee accounts for approximately 70% of the overall costs, it is common practice in the industry to benchmark efficiency and practices computing unit costs for kg of parchment produced.

The survey we implemented in the field, conducted through face-to-face interviews with managers at the stations by people with significant experience in the industry, therefore focused on understanding unit costs and their various components. By working together with the managers through the stations accounts for the season, it was possible to obtain what we believe are accurate measure of unit costs and their breakdown across a variety of components. The unit costs of mill i can be decompose as follows,

$$UC_i = \left(P_i^{kg} \times CR_i\right) + OC_i \tag{1}$$

where UC_i are the unit costs, P_i^{kg} is the price per kilogram of cherries paid by the mill, CR_i is the conversion ratio and OC_i are other costs, mainly labour, finance and transport.[4]

Figure 3 and Table 2 document dispersion in the unit costs of the mills and their components. Figure 3 shows that a significant proportion of the observed dispersion in unit costs can be explained by differences in the geographic characteristics of the area in which the coop operates (historical suitability for coffee, availability of trees, elevation, slope, etc.) as well as capacity installed. However, even purging the data from differences in costs driven by these factors, Figure 3 still documents significant dispersion in unit costs. In particular, within this narrowly defined sector with essentially identical production processes across firms and accounting for both geographical and technological differences, we still find the 90/10 ratio of unit costs to be 1.5.

Table 2 decomposes this dispersion in unit costs into its components. As expected, there is essentially no dispersion in the conversion ratio CR_i (90/10 ratio is lower than 1.1). There is moderately more dispersion in the cherry prices paid by the stations to farmers (90/10 ratio is 1.32). Together, the conversion ratio and the price of cherries account for approximately 70% of the unit costs. It follows that the bulk of the dispersion in unit costs originates from dispersion in the other components of unit costs which are directly under the control of management. Here we find a 90/10 ratio equal to 2.32. The remaining of the analysis studies the impact of competition on this efficiency dispersion.

[3]The exact conversion ratio might be affected by the exact coffee variety grown in the area and other geographical factors affecting the organic properties of coffee.

[4]Below we analyze the breakdown of these costs as well.

5 Theory

This section lays out a simple theoretical framework to discipline the empirical analysis.

5.1 Set-Up

We consider a coffee producing region which, for simplicity, we represent as a circle of radius r. At each location inside and on the circle there is a mass 1 of farmers, with supply function $c(q) = \frac{q^{\gamma+1}}{\gamma+1}$, $\gamma > 0$. To process coffee, the stations has to enter the market installing capacity at cost $K(Q)$. Having paid these fixed costs, the station can process Q units of inputs during the harvest season. The unit processing cost is denoted with $c(N)$. For simplicity, this unit cost includes both (working) capital and labour costs. It is assumed to be constant, provided the firm operates below capacity (and becomes infinite for production above capacity). In a reduced form, the unit cost is allowed to depend on the number of competitors, N. There is a unit transport costs, τ, per unit of produce and distance. That is, we assume it costs an amount $T(q, d) = \tau dq$ to transport q units from distance d.[5] The final price is given by v.

The timing of events is as follows. First, the firm decided whether and, if yes, where to enter. Then the firms decides in which local markets to buy cherries, taking as given the choice of its opponents. Finally, prices and purchases in the local markets are realized and production takes place. Payoffs are realized. The model is solved in steps, by backward induction. The market local equilibrium is derived given a number of competing firms.

5.2 Equilibrium

Step 1: *Definition of Local Market Equilibrium*

Suppose there are n_d firms purchasing in location d (from now on, omit the subscript if it doesn't create confusion). Assume there is a continum of small, price taker, farmers i. Given price $p(n, \mathbf{D})$ the aggregate quantity supplied is given by $\int_0^1 [p(n, \mathbf{D})]^{\frac{1}{\gamma}} di$, i.e., $(q_d + q_{d'})^\gamma = p$. This quantity is shared across all firms in the market, but not equally, since they face different marginal transportation costs. We assume that each firm at distance d gets a share $\sigma(d, \mathbf{D}_{-d})$ of the resulting aggregate quantity, with $\sigma_d < 0$, and $\sigma_{d'} > 0$. Similarly, we assume that $p_n > 0$, $p_d < 0$.

[5] An extension allows for fixed costs of transport, which generate additional extensive margins of response to competition, as well as variable capacity.

Cournot Example. Suppose firms compete cournot, i.e., they announce quantity they want to purchase in the location. For simplicity, assume there are only two firms. The price to be paid in the market is equal to the marginal costs of producing the aggregate quantity, i.e., $p = (q_d + q_{d'})^{\gamma}$. Assuming the firm hasn't reached the capacity constraint, the decision of how much to source is separable across location. The firms sets q_d to maximize profits, i.e.,

$$\max_{q_d} (v - c - \tau d - p(q_d, q_{d'}))\, q_d \tag{2}$$

which gives the first order condition

$$(v - c - \tau d - p(q_d, q_{d'})) - \gamma \frac{(q_d + q_{d'})^{\gamma}}{q_d + q_{d'}} q_d = 0. \tag{3}$$

Substituting for the level of prices in the locality implied by the supply function, the previous expression becomes

$$(v - c - \tau d - (q_d + q_{d'})^{\gamma}) = \gamma \frac{(q_d + q_{d'})^{\gamma}}{q_d + q_{d'}} q_d$$

In the special case in which the supply function is linear, i.e., $\gamma = 1$, we obtain an explicit solution linking the quantity purchased by the station in locality d as a function of unit costs, transport costs and the quantity purchased by the competing firm,

$$q_d = \frac{(v - c - \tau d - q_{d'})}{2}. \tag{4}$$

By symmetry, the (unique) Nash equilibrium is given by

$$q_d^* = \frac{1}{3}(v - c - \tau d) - \frac{1}{3}\tau \left(d - d'\right) \tag{5}$$

with corresponding price received by farmers in locality d given by

$$P^* = \frac{2}{3}(v - c) - \frac{\tau}{3}\left(d + d'\right) \tag{6}$$

Farmers in localities which are closer to stations (lower average d and d') and served by better roads (lower τ), therefore, receive better prices. The logic of the previous expression can be generalized to the case in which N firms are active in locality d, yielding

$$P^* = \frac{N}{N+1}(v - c) - \frac{\tau}{N+1}\left(\sum_{n \in N} d_n\right).$$

In contrast, a firm that operates as a monopolist in locality d, sets

$$q_d^M = \frac{(v - c - \tau d)}{2}$$

and pay prices to farmers equal to

$$P_m^* = \frac{(v - c - \tau d)}{2}.$$

So, as a monopolist a firm purchases up to localities at distance $d_m : v - c - \tau d - \frac{(v-c-\tau d)}{2} = 0$, i.e., until $d_m = \frac{v-c}{\tau}$. As a duopolist, instead, the firm purchases cherries up to localities at distance $d_D = \frac{(v-c-\tau d')}{4\tau}$.

Step 2: *Entry Decision*

Taking as given the decision of other firms and anticipating the realized prices and quantities in each locality, each firm decides whether to enter in a particular location d. We assume that each entry decision is independent. Let us consider symmetrical equilibria (ignoring boundaries). The equilibrium on the circle can be found as follows:

1. - suppose there are N stations equally spaced.

   - assume stations have a boundary around themselves such that their competitors do not source inside those boundaries. The boundary is defined by $d' = \frac{(v-c-\tau d)}{4}$, where $D = d + d'$ defines the total distance between the two stations.

   - Question 1: does a firm have incentives to travel to the mid-point, if it was a monopolist? YES iff $\frac{v-c}{\tau} > \frac{D}{2}$.

   - Question 2: does a firm have incentives to be there as a duopolist? YES iff $\frac{\left(v-c-\tau\frac{D}{2}\right)}{4\tau} > \frac{D}{2}$.

2. Then, *assume* that close to the station there the competitor as no incentive to come, i.e., $\frac{(v-c)}{4} < D$.

3. Under assumption 2, equilibrium profits, quantities and prices can be characterized.

5.3 Empirical Predictions

The model delivers predictions for the following variables:

1. **Entry**: as a function of τ (density of roads), c_f (the slope of the supply function), and suitability s (to be parametrized) the model gives you the "density" of firms and how much they compete;

2. **Prices**: determinants of local prices and average prices paid by stations;

3. **Test for Pecuniary Externality**: prices might not increase with competition, unit costs increase, if unit costs are increasing in N, $c'(N) > 0$.

[TO BE COMPLETED]

6 Empirical Results

6.1 Efficiency and Competition: OLS Estimates

[TABLE 3 HERE]

[FIGURE 5 HERE]

6.2 Entry Model and Instrument Construction

Estimating the causal impact of entry and competition on market participants poses severe methodological challenges. First, the entry decision of a wet mill is based on expected profits considerations. Many of the determinants of expected profits that are typically not observed by the econometrician (for instance, the potential to produce high quality cherries) are likely to correlate with outcome of interests (e.g., efficiency in unit costs, farmer input decisions). Second, conditional on entry, the causal impact of further competition on outcomes of interest is also challenging to identify. For example, the presence of a competitor near a well-managed wet mill is consistent with both competition improving management, or with better managers selecting locations that have higher potential for profit and, therefore, attract higher entry in those areas. Below we first describe how we construct a model that predicts entry, which is based on a our detailed knowledge of the placement of wet mills and than we describe our empirical identification strategy to establish causality, via. instrumental variables.

The standard approach in industrial organization is to structurally estimate a particular entry model specified ex-ante by the researcher (see, e.g., Barry and Reiss (2007) for a review of the literature). The methodology proposed in this paper, instead, relies on a reduced form approach that exploits the spatial nature of the industry alongside very detailed knowledge of the entry process used by investors in the sector. Wet mills cannot be located everywhere: they need to be close to coffee producing areas and require easy access to roads and water. In the early years of the industry, a program coordinated by USAID produced a engineering model for the optimal placement of wet mills in Rwanda (see, USAID 2006). The model was then used by the regulating bodies and by private investors to provide possible areas for the placement of wet mills. The model recommends sites that are suitable for the placement of wet mills based on a vector of characteristics which are then aggregated into a score. The score summarizes how appropriate a particular location defined at a very high spatial resolution is suitable for the establishment of a wet mill. We re-construct the individual components of the engineering model and also take into account the appropriate Government of

Rwanda legislation on the licensing and construction of coffee wet mills:[6]

Criteria #1: Select all sectors with tree count > 30,000. Wet mill's need a constant supply of coffee cherries and hence should be placed near major coffee-growing areas. Coffee trees can take up to 3 years before they can produce cherries, hence it is important for stations to be placed near the appropriate mature coffee trees, usually trees from 3 to 30 years old are considered as productive.[7] In Figure 3 we illustrates that all wet mills enter in sectors with at least 30,000 coffee trees (threshold depicted by red line at 10.31).

[FIGURE 4 HERE]

Criteria #2: Potential site's for wet mill's cannot be further than 3 km's from a spring source. Wet mill's require a source of clean and plentiful water nearby the wet mill, this is primarily to wash the cherries, distance's more than 3 km's are in most instance uneconomical for wet mills as piping water over large distance on uneven terrain is expensive. Using a detailed Geographic Information System (GIS) dataset we are able to map out all the rivers of different hierarchies in Rwanda, in particular our interest is on the small rivers which form the end points of springs. From a master river database we first isolate the smallest rivers and obtain their origins by using the Shutter Radar Topography Mission raster data to obtain elevation of origins at 10 m contours to ascertain the origins of rivers and hence end points of springs. We than overlay this on a GIS map of Rwanda that splices the country into one square kilometer grid boxes (henceforth, grid). This allows us to determine whether a grid has a spring and given the geo-spatial nature of our data it also allows us to determine the distance to the closet spring. Table 4 column 1 and 2 show that the presence of a spring in a grid, indicated by a dummy which takes a value of 1 if the grid has a spring, predicts the entry of wet mills well, conditioning on satisfying criterion 1, the presence of sufficient trees in the sector. In column 1 a simple probit estimation is run without controlling for any geographical characteristics which could potentially determine the profitability of entry. In column 2 we include additional geographical characteristics both as polynomials and including their cross interactions.

[6]The National Export Board (NAEB) has put legislation in place that establishes the criteria on the necessary requirements to obtain a license to construct a coffee washing station, see Ministry instructions for construction of coffee washing stations in Rwanda Ref No. 041/1500/MIN/06, official title *Amabwiriza Ya Minisitiri Yerekeye Iyubakwa Ry'inganda Zitunganya Kawa.*

[7]Rwanda is administratively divided into provinces, which are further divided into districts, in turn into sectors, which are than divided into cells and these are divided into villages.

These geographical characteristics (all computed at the 1 square kilometer grid) are elevation, slope, historical suitability of the area for coffee growing, length of river, number of coffee trees, and longitude and latitude. Various other definitions of the presence of the river are used to check the robustness of the results, these include a simple dummy whether a small river is within 3 km of that particular grid, or wether a spring source is within 2 kilometers, results are stronger the stringent the condition.

Criteria #3: Potential site's for wet mill's need access to roads to transport the cherries once they are processed, roads should be at least within 1 kilometer of a site. Cherries are transported often by women on foot or bicycles carrying up to 30 kgs in plastic buckets on their heads. The primary need for the access to roads is to transport the dried parchment to Kigali for milling and exporting. We use a GIS layer of the most detailed roads inventory available that identifies up to minor tracks from satellite orthophotos (circa 2008/09). Due to the spatial nature of the data we can not only determine the length of roads (kms) in each grid but also the distance from a particular grid to the nearest road of any type. Column 3 and 4 shows that the presence of roads within 1 km of a grid location predicts entry of a wet mill, indicated by a dummy that takes a value of 1 if the grid has within 1 km of its location a unpaved local road, conditional on the sector having sufficient coffee trees. Column 3 is the unconditional coefficient without controlling for any geographical characteristics which could determine entry of a wet mill. Column 4 includes geographical characteristics (polynomials and interactions).

Note the availability of both a nearby water source and transportation predict's entry conditional on the sector in the grid having access to sufficient coffee trees. In column 5 we include both the presence of spring within a 3 km radius of a potential site and the same grid location having access to a road within a 1 km radius. Column 6 includes our set of geographical controls their polynomials and interactions. In column 7 an interaction is attempted between water source and the presence of roads, which shows us that the presence of roads and springs are substitutes conditionally on wet mill entry in the first place.

Score

Each of the these criteria can be used to create a *score*, a numerical value for each grid in Rwanda to determine how suitable a grid is for the location of a wet mill. The steps are as follows, our probit entry model is run using the criteria if the grid has a spring source and access to roads, conditional on the grid box being present in a sector with sufficient coffee trees and the land cover being suitable for constructing a wet mill (i.e.

grids, which are in built-up areas, on water bodies, on runways, in natural forest are excluded). Predicted value of the score are computed post estimation of this model, which are displayed in Figure 7. This than allows us to compute for each grid a score metric that our model has predicted, note in order to do this first we had to validate our entry model.

[FIGURE 7 HERE]

[TABLE 4 HERE]

6.3 Efficiency and Competition: IV Estimates

We are interested in investigating the causal impact of competition in the procurement of cherries on a variety of the wet mill's outcomes, including unit costs, conversion ratios, utilization rates, managerial practices, etc. As mentioned this is empirically challenging to provide evidence on the effect of competition on efficiency, why? Precise measurement of both productivity and competition is rare in low income countries, we need across firms within a sector. Second, entry of competitors is likely to be a result as much as a cause of firm's poor management/performance.

We measure competition as the number of competing mills reported by the mill's manager to be purchasing within the catchment area of the mill. Given the geospatial nature of our data we are able to for each grid that contains a wet mill, obtain the distance to the nearest wet mill and the distance to the second nearest wet mill.

Let's denote y_i the outcome of interest and by C_i competition for mill i. We are interested in estimating

$$y_i = \alpha + \beta C_i + \gamma_0 X_i + \gamma Z_i + \gamma^N Z_i^N + \varepsilon_i$$

where X_i, Z_i and Z_i^N are comprehensive vectors of controls at the mill level (i) and ε_i is an error term. The vector X_i includes controls for capacity installed, age and type (cooperative, private and NGO supported) of the mill. The vector Z_i includes all potential drivers of the performance of mills, including geographic and soil characteristics both at the location of the mill (within a radius of 0-5 kms) as well as within a radius of 5-10 kilometers of the wet mill (according to out field work wet mills think of competition for cherries up to 10 km from their site as the potential catchment area). These controls are polynomials in (i) suitability for coffee, (ii) elevation and slope, (iv) presence and types of roads, (v) presence and type of rivers, (vi) density of coffee trees,

and vii) interactions of these controls. The vector Z_i^N includes the (weighted) average of the same variables in the area surrounding the mill's catchment area (in a radius of 5 to 10 kilometers).

In this specification, competition is likely to be correlated with the error term. For instance, competing stations might source coffee more aggressively and penetrate the catchment areas of mills that are relatively poorly run. For this reason, we instrument for competition using the (predicted) score from the engineering model in the areas around the catchment area of the mill, i.e the predicted score of the area between 5-10 km. The score is a predicted value obtained after estimating the entry model as described above, in particular column 6 in Table 4.

The predicted score predicts the number of wet mills around the neighboring grid within the 5-10 km radius. Specifically, the first stage is given by

$$C_i = \alpha + \widehat{\beta} S_i^N + \widehat{\gamma}_0 X_i + \widehat{\gamma} Z_i + \widehat{\gamma}^N Z_i^N + \mu_i$$

where S_i^N is a weighted average of the predicted score in the neighbors of mill's i catchment area (5-10 km). In other word, the exclusion restriction is satisfied if, conditional on all costs drivers locally, i.e. at and around the mill location as well as in its neighborhood, drivers of placement of mills in the neighbors are uncorrelated with determinants of mill's outcomes other than competition.

[TABLE 5 HERE]

[FIGURE 6 HERE]

[TABLE 6 HERE]

7 Mechanisms and Other Outcomes

7.1 Capacity Utilization

[TO BE ADDED]

7.2 Interlinked Transactions

[TO BE ADDED]

7.3 Farmer's Outcomes

[TO BE ADDED]

7.4 Trust

[TO BE ADDED]

8 Conclusion

[TO BE ADDED]

References

[1] Berry, S. and P. Reiss (2007) Empirical Models of entry and market structure. Armstrong M, Porter R, eds. *Handbook of Industrial Organization*, Vol. 3 (North-Holland, Amsterdam), 1845-1886.

[2] Bloom, N. and J. Van Reenen (2007) "Measuring and Explaining Management Practices across Firms and Countries", *Quarterly Journal of Economics*, 122(4), pp. 1351-1408.

[3] Bloom, N., B. Eifert, A. Mahajan, D. McKenzie and J. Roberts (2013) "Does Management Matter? Evidence from India", *Quarterly Journal of Economics*, 128(1), pp. 1-51.

[4] Blouin, A. (2014) "Culture and Contracts: The Historical Legacy of Forced Labour", *mimeo.*

[5] Blouin, A. and R. Macchiavello (2013) "Tropical Lending: International Prices, Strategic Default and Credit Constraints among Coffee Washing Stations" *mimeo*

[6] Binswanger, H. P. and M. R. Rosenzweig (1986) "Behavioural and material determinants of production relations in agriculture", *The Journal of Development Studies*, 22(3), pp. 503-539.

[7] Diamond, P. (1987) "Consumer differences and prices in a search model", *Quarterly Journal of Economics*, 429-436.

[8] Fafchamps, M. (2004) *Market Institutions in Sub-Saharan Africa*, MIT Press, Cambridgem Mass.

[9] Ghosh, P. and D. Ray (1996) "Cooperation in Community Interactions without Information Flows", *Review of Economic Studies*, 63: 491-519.

[10] Hosios, A. (1990) "On the efficiency of matching and related models of search and unemployment." *The Review of Economic Studies* 57.2: 279-298.

[11] Hsieh, C. and P. J. Klenow (2009) "Misallocation and Manufacturing TFP in China and India", *Quarterly Journal of Economics*, 124(4), pp. 1403-48.

[12] Hsieh, C. and B. Olken (2014) "The Missing 'Missing Middle'", *Journal of Economic Perspective*, forthcoming

[13] Hsieh, C. and E. Moretti (2003) "Can Free Entry Be Inefficient? Fixed Commissions and Social Waste in the Real Estate Industry", *Journal of Political Economy* 111-5.

[14] Kranton, R. (1996) "Reciprocal exchange: a self-sustaining system." *American Economic Review*, 830-851.

[15] Mankiw, N. and M. D. Whinston (1986) "Free Entry and Social Inefficiency", *The RAND Journal of Economics*, 17(1), pp. 48-58.

[16] Mullainathan, S. and S. Sukhtankar (2014) "Ownership Structure and Economic Outcomes: The Case of Sugar Mills in India", *mimeo*

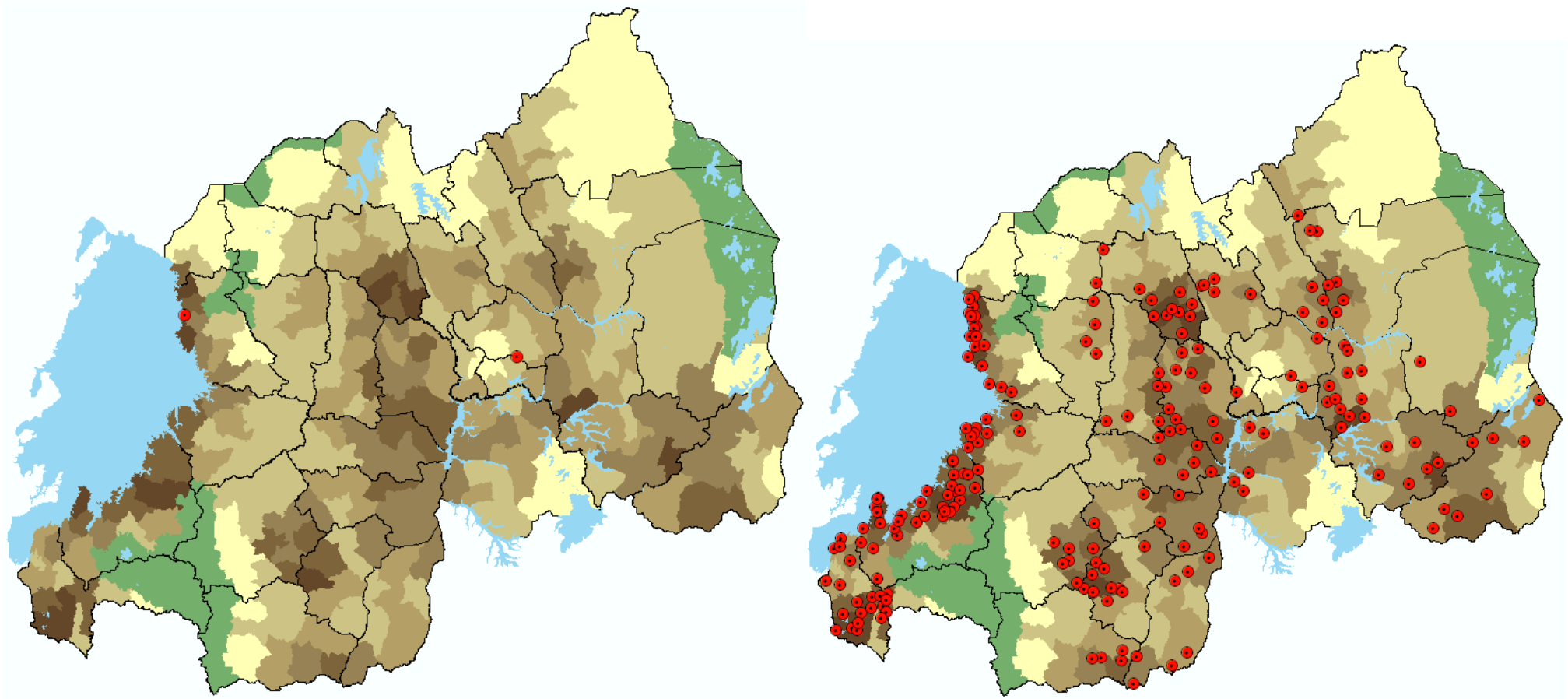
[17] Söderbom, M. and S. Kamarudeen (2013) "Constraints and opportunities in Rwanda's industrial sector", International Growth Centre

[18] Syverson, C. (2004) "Product Substitutability and Productivity Dispersion", *Review of Economics and Statistics*, 86(2), pp. 534-50.

[19] Syverson, C. (2011) "What Determines Productivity?", *Journal of Economic Literature*, 49(2), pp. 326-365.

[20] World Bank (2008) "World Development Report: Agriculture for Development", World Bank

FIGURE 1: WET MILL ENTRY IN RWANDA

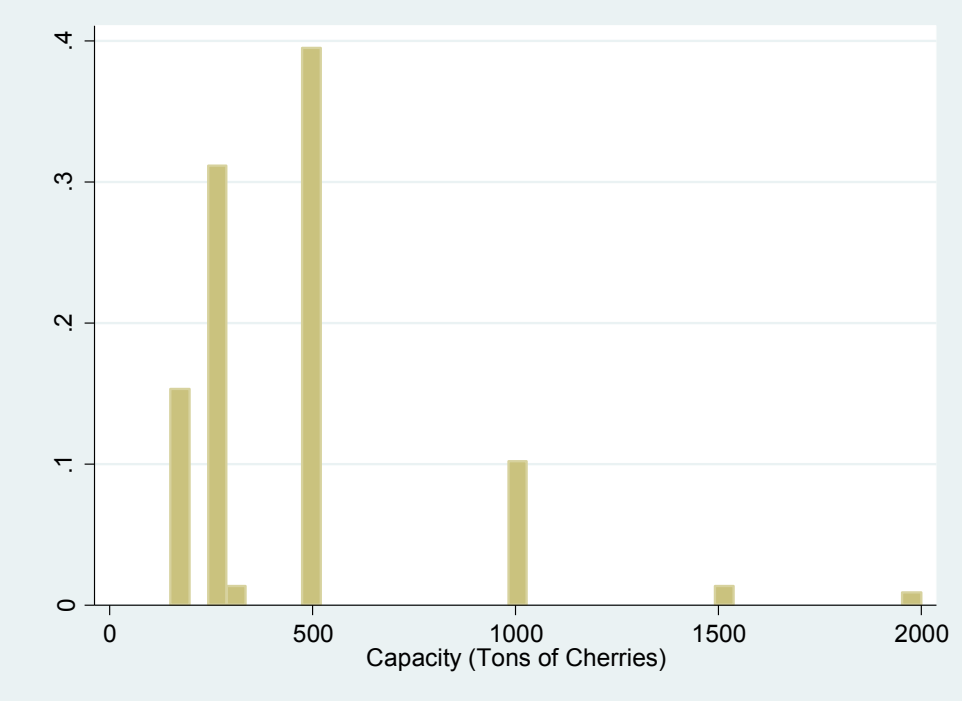


2002 **2012**

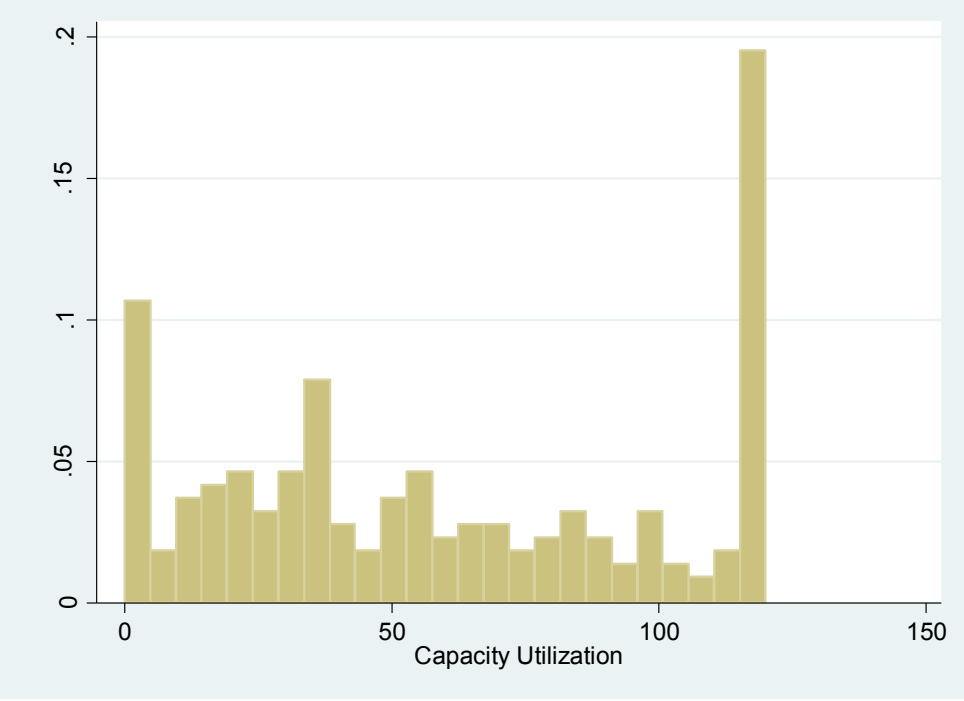
Notes: This figure illustrates, the spatial and temporal evolution of the entry of wet mills in Rwanda. In 2002 there were only 2 wet mills in Rwanda (denoted by the red dots). By 2012 there has been entry of up to 210 wet mills across Rwanda. Green shades indicate national parks, blue indicates water bodies, the background is number of coffee trees at the sector level, the darker the color the higher the number of trees in the sector. Source: Authors calculation and GIS data from various Government agencies.

FIGURE 2: INSTALLED CAPACITY AND UTILIZATION, 2012 SEASON

Panel A.

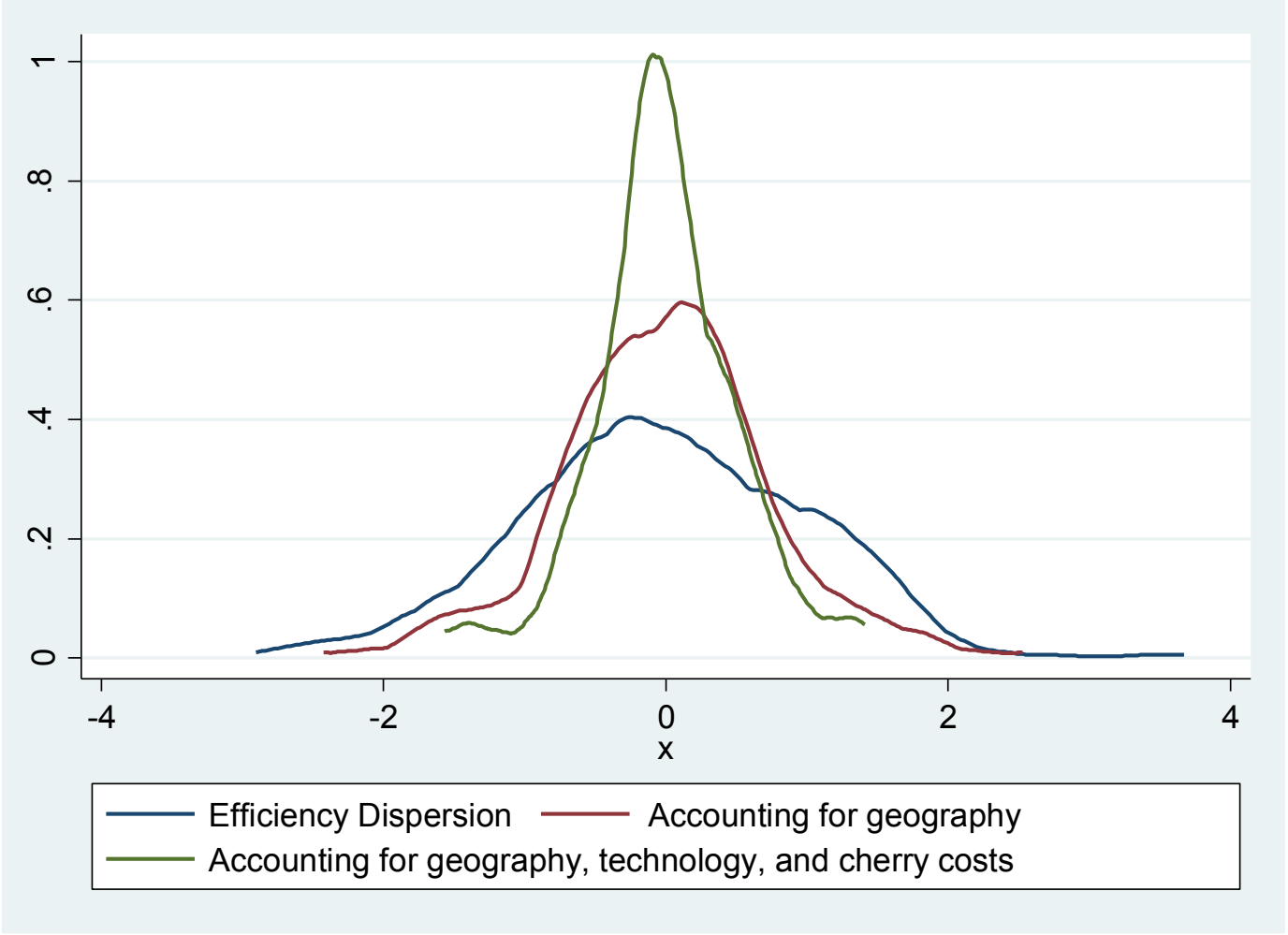


Panel B.



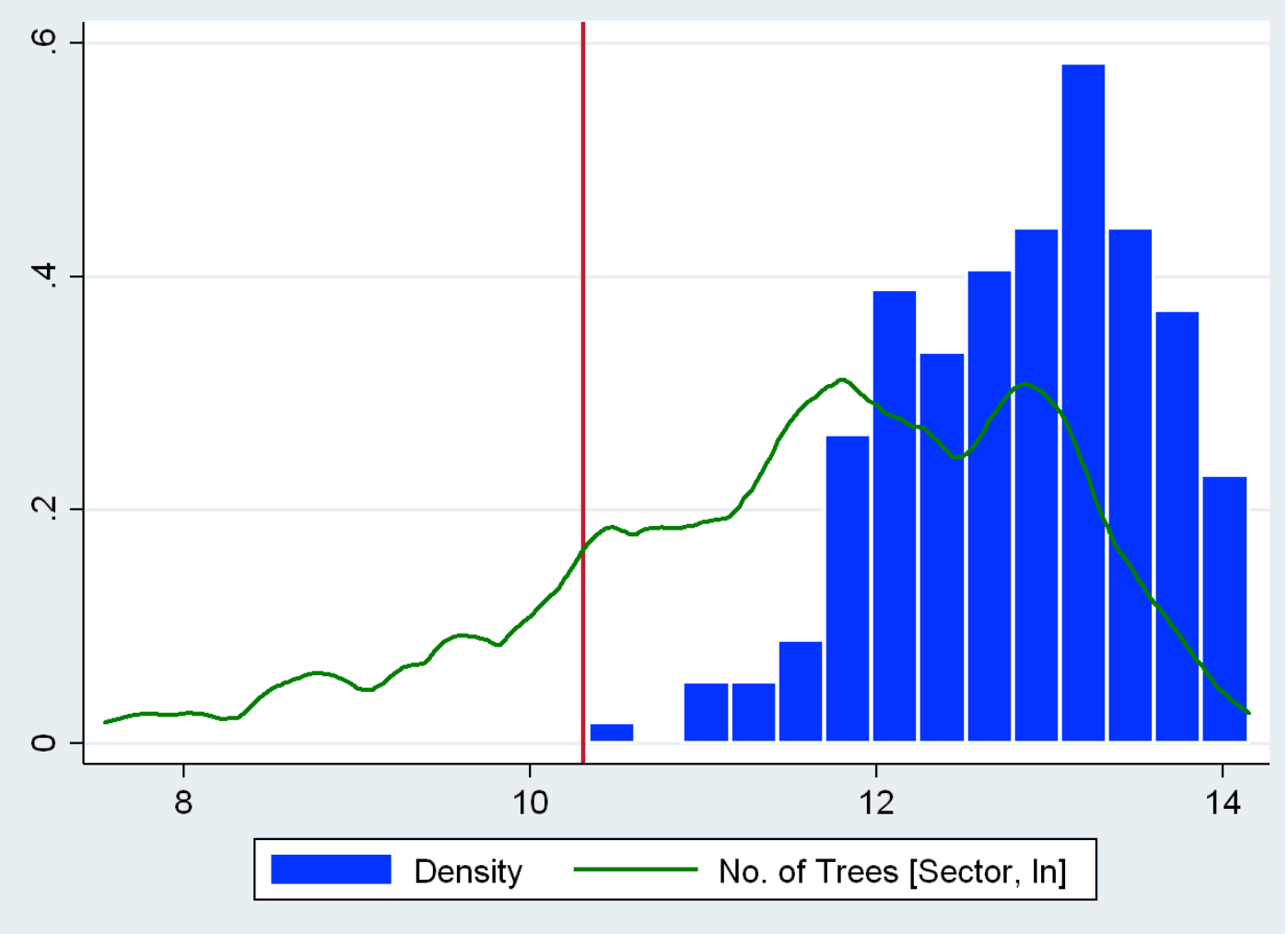
Notes: Panel A. displays the distribution of capacity installed by wet mills. The figure shows significant dispersion in capacity installed with many small stations. Panel B. displays the distribution of capacity utilization by stations for the 2012 season. In aggregate only 60% of the installed capacity as utilized. In the survey approximately 40% of the stations reported to have had days in which they wanted to purchase cherries but could not. Source: Author's calculation on NAEB administrative data.

FIGURE 3: DISPERSION IN EFFICIENCY



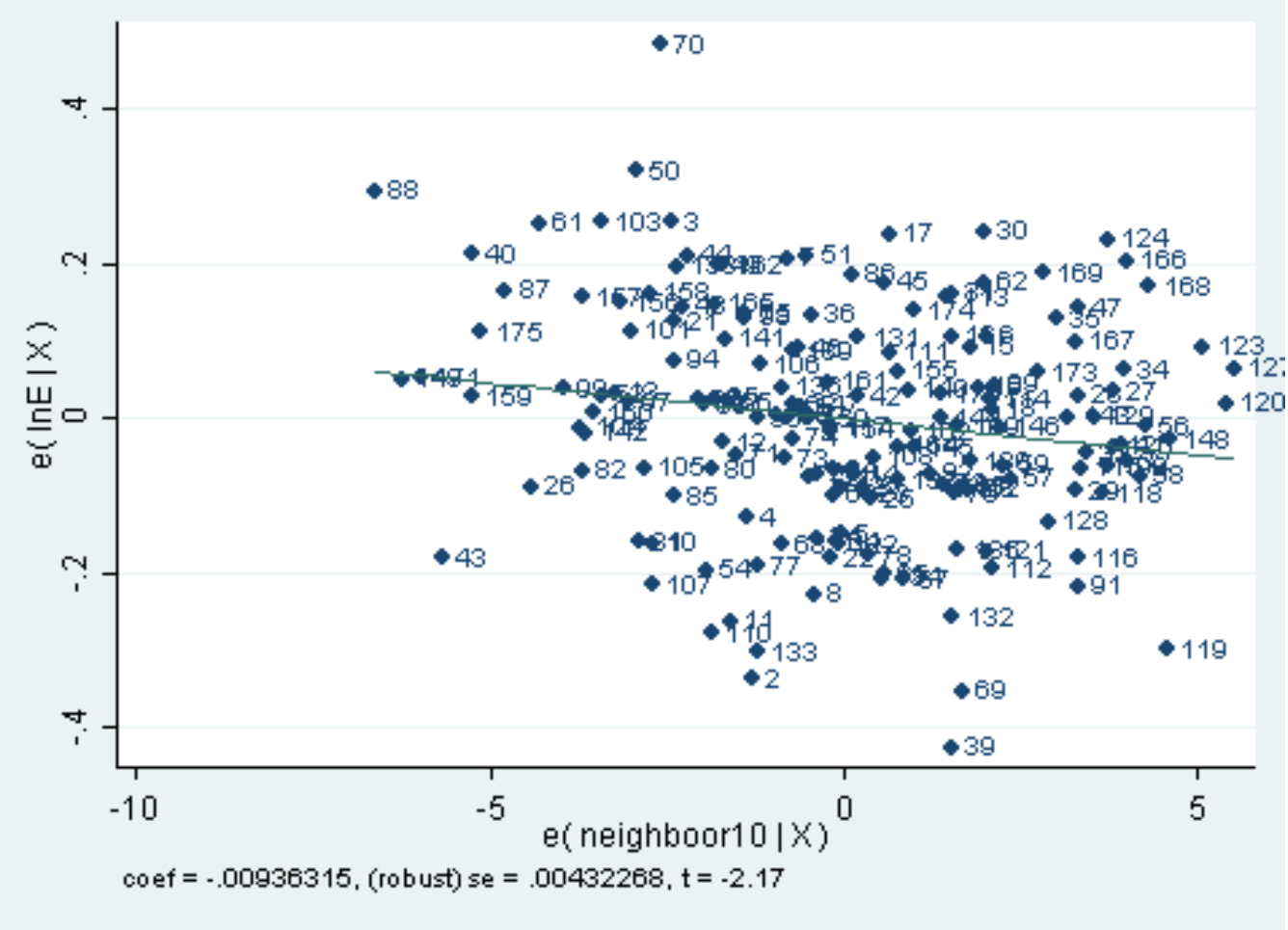
Notes: This figure shows the distribution (standardized) of efficiency, which we define as the total unit cost of obtaining one kilogram of parchment (in blue). We then control for local geography (in red) of the wet mill (elevation, slope, historical suitability of coffee, density of coffee trees in the sector, length of road and river and area) and the type of wet mill (private or cooperative) and the local cost of purchasing cherries (in green). Source: Author's calculation on Wet Mill Survey (2012) and GIS data.

FIGURE 4: ENTRY OF WET MILL AND PRESENCE OF COFFEE TREES



Notes: This figure plots the kernel density of coffee trees in all administrative sectors of Rwanda (in green) and overlays as a histogram, that shows the number of coffee trees in the sector in which a wet mill has entered (in blue). The red line indicates the threshold the engineering model recommends the minimum number of trees that should be available in a sector to attract a wet mill. *Source*: Author's calculation on Coffee Census (2009) and Wet Mill Survey 2012.

FIGURE 5: EFFICIENCY OF WET MILL AND THE NUMBER OF WET MILLS IN THE NEIGHBOORHOOD (5-10 KM)



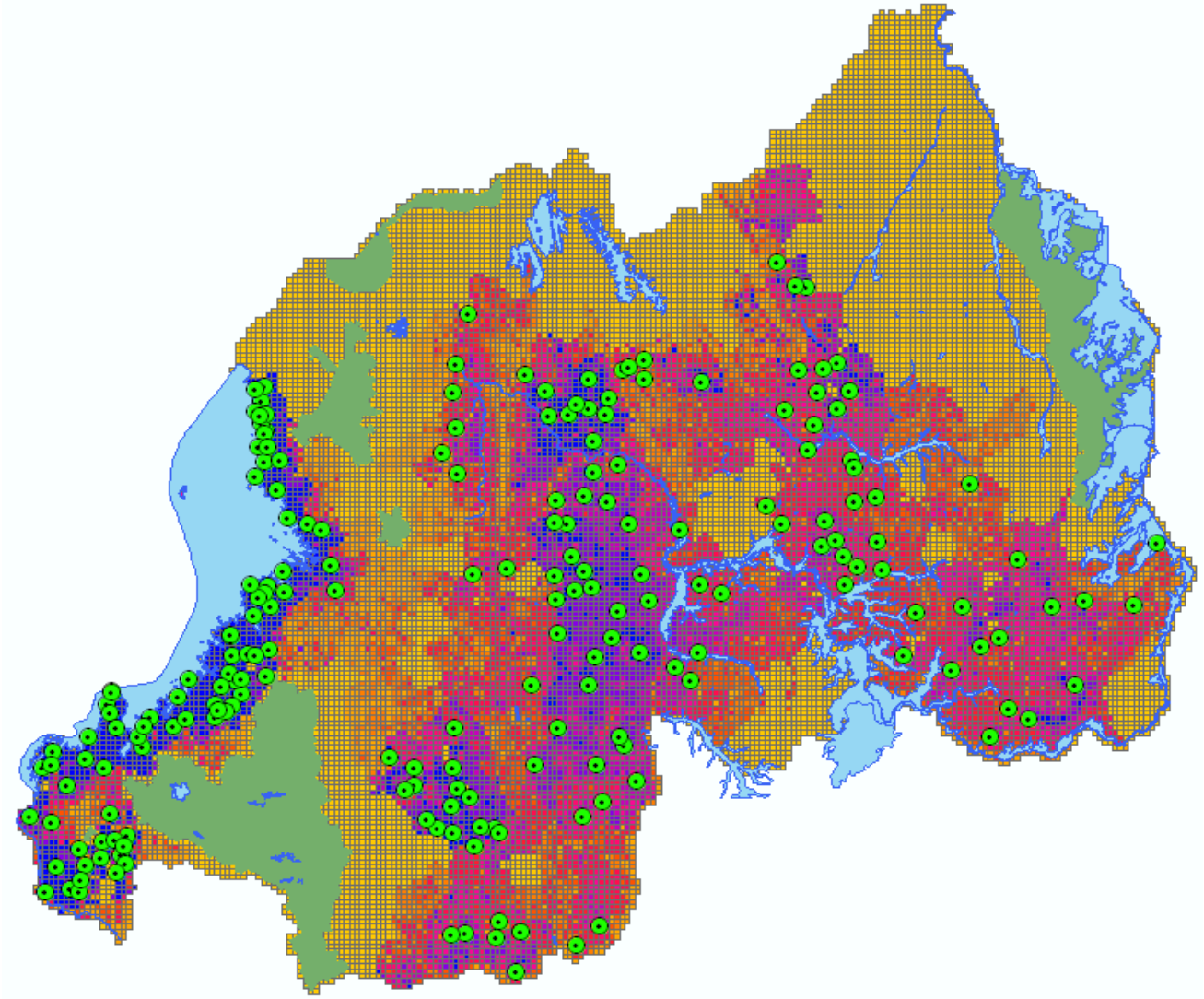
Notes: This figure plots the OLS estimation of efficiency (ln) on the number of coffee wet mills in the radius of 5-10 km around the wet mill. The regression includes the following controls: entry model criteria (dummy variable indication presence of river and road within 5 km of the wet mill), geographical controls (elevation, slope, soil suitability, length of river, longitude and latitude) and wet mill controls (type of wet mill, installed capacity, Technoserve support, age of station and its square).

FIGURE 6: EFFICIENCY OF WET MILL AND THE PREDICTED *SCORE* OF NUMBER OF WET MILLS IN THE NEIGHBOORHOOD (5-10 KM)



Notes: This figure plots the reduced form result of the OLS estimation of efficiency (ln) on the suitability score at the radius of 5-10 km around the wet mill. The regression includes the following controls: entry model criteria (dummy variable indication presence of river and road within 5 km of the wet mill), geographical controls (elevation, slope, soil suitability, length of river, longitude and latitude) and wet mill controls (type of wet mill, installed capacity, Technoserve support, age of station and its square).

FIGURE 7: PREDICTED SCORE FOR OPTIMAL PLACEMENT OF WET MILL



Notes: This figure illustrates, the predicted "score" for the placement of a wet mill in each grid box (i.e. 1 square kilometre) in Rwanda using our model of entry, which is driven by engineering considerations of the optimal placement of wet mills. The darker the color (the color ramp is going from yellow-pink-to blue) the higher the probability of entry occurring. Green dots illustrate existing wet mills and green blocks depict national parks and pale blue indicates water bodies.

TABLE 1: DESCRIPTIVE STATISTICS

Variable:	Mean	Median	Private vs Cooperatives	Non-*TNS* vs. *TNS* Cooperatives
Theoretical Capacity (Tons)	447	500	162.59***	307.02***
End of Season Utlization Rate	47.20%	32.40%	-11.36***	-14.47***
Average Price Paid to Farmers (RWF)	208.53	200.00	4.44***	19.31***
Output per worker	3.98	3.11	3.4**	1.06**
Workers per ton of capacity installed	0.21	0.18	-0.04**	-0.11**
Permanent supervisors	0.17	0.00	0.09	-0.21***
Seasonal supervisors	2.31	2.00	0.48**	1.26***
Permanent collectors	0.36	0.00	0.20	0.43
Seasonal collectors	17.98	12.00	0.08	7.27**
Production of fully-washed green coffee this season (Tons)	46.01	32.00	18.09116***	15.64564**
Cherries purchased to date (Tons)	294.82	199.91	120.4038***	132.4915***
Cherries washed to date (Tons)	60.21	40.00	28.56917***	25.82388**
Fully-washed parchment to date (Tons)	43.91	30.00	19.78704***	19.42955***
Unit Costs [RWF per Kg]	1792.99	1800.00	97.93***	149.80**
Altitude (m)	1584.16	1570.50	1.615	4.169
Historic Suitability for Coffee	0.50	0.485	0.002	-0.021*
Current Suitability for Coffee	0.61	0.606	-0.014	-0.052
Distance to nearest paved road (km)	14.52	10.00	1.031	4.466
Months Manager Employed	8.60	12.00	2.611***	3.765***
Manager's Income from Station (in $)	1621	1187	1340***	824***

Notes: The above table illustrates various characteristics across the different types of wet mill's (private, cooperatves and Technoserve supported). Column 3, compute's t-statistics testing for difference in means between private versus cooperative wet mills. Column 4 conducts test within cooperatives between Technoserve supported cooperative wet mills versus non-Technoserve coopertives. ***, **, * denote statistically significance

TABLE 2: DISPERSION IN EFFICIENCY

Variable:	**Unit Costs**	**Physical Efficiency**	**Cherry Prices**	**Other Unit Costs**
p75 - p25 Ratio	1.22	1.04	1.11	1.66
p90 - p10 Ratio	1.51	1.1	1.32	2.34
% Variation explained by:				
Geography	27	34	41	30
Technology	8	6	7	4
Physical Efficiency	10	--	[4]	[3]
Cherry Prices	5	[4]	--	[7]
% Variation Unexplained:	**50**	**60**	**52**	**57**

Notes: For Cherry costs: R75/25=1.15; R90/10= 1.28; above we tabulate the dispersion in terms of comparing the top 10 percentile with the bottom 10 percintle as well as the top 25 percentile with the bottom 25 percentile. The unit cost is the cost (in RfW) in processing to obtain 1 kilogram of parchment. Unit cost can be broken down into its various components - the physical efficiency (conversion of x kgs of cheeries for 1 kg of parchment, the cost of purchasing cherries and other unit costs (labor, capital, procurement and transport).

TABLE 3: OLS - A FIRST LOOK AT COMPETITION AND WET MILL OUTCOMES

	Panel A					Panel B				
Dependent Variable:	**Ln (Efficiency)**					**Ln (Other Costs)**				
Sample	**Surveyed Wet Mills only**					**Surveyed Wet Mills only**				
	[1]	[2]	[3]	[4]	[5]	[1]	[2]	[3]	[4]	[5]
No. of Wet Mills within 5-10 km	-0.0094**		-0.011**			0.004		0.015		
	(0.004)		(0.005)			(0.011)		(0.012)		
No. of Wet Mills within 5 km		-0.0074	0.0049				-0.0169	-0.034		
		(0.008)	(0.0097)				(0.023)	(0.025)		
Number of Wet Mills sourcing from same collection site (survey question)				-0.046***	-0.039**				0.0781**	0.065*
				(0.017)	(0.018)				(0.038)	(0.039)
Number of other Buyers sourcing from same collection site (survey question)					-0.02**					0.043**
					(0.008)					(0.019)
Controls										
Entry model criteria	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Geographical characteristics	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Wet Mill characteristics	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
R2	0.174	0.154	0.175	0.203	0.225	0.1469	0.1505	0.1583	0.2021	0.2254
Number of observations	171	171	171	164	163	170	170	170	164	163

	Panel C					Panel D				
Dependent Variable:	**Ln (Price)**					**Ln (Conversion Ratio)**				
Sample	**Surveyed Wet Mills only**					**Surveyed Wet Mills only**				
	[1]	[2]	[3]	[4]	[5]	[1]	[2]	[3]	[4]	[5]
No. of Wet Mills within 5-10 km	0.008**		0.004			0.0005		0.0029		
	(0.004)		(0.005)			(0.0014)		0.001		
No. of Wet Mills within 5 km		0.0158**	0.011				-0.004	-0.007*		
		(0.007)	(0.009)				(0.003)	(0.004)		
Number of Wet Mills sourcing from same collection site (from survey)				0.035**	0.034**				-0.009	-0.012*
				(0.015)	(0.015)				(0.006)	(0.006)
Number of other Buyers sourcing from same collection site (from survey)					0.003					0.006**
					(0.006)					(0.002)
Controls										
Entry model criteria	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Geographical characteristics	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Wet Mill characteristics	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
R2	0.3182	0.3205	0.3249	0.3292	0.3298	0.096	0.107	0.3249	0.3292	0.1293
Number of observations	158	158	158	155	154	159	159	158	155	154

Notes: Standard errors in ( ) are clustered at the sector level (170). ***, **, * denote statistically significance at 1, 5, 10 percent, respectively. Outcomes of interest change across the different panels. Entry model controls are the dummy variables that indicated presence of spring and road both within 5 km of the wet mill, geographical controls include the elevation, slope, soil suitability, length of river within 5 km of wet mill and wet mill controls are a dummy if Technoserve supported station, the type of wet mill (private or cooperative), the age (and age square) and installed capacity of the wet mill.

TABLE 4: ENTRY MODEL OF WET MILL'S

	Spring Criterion		**Roads Criterion**		**Spring and Roads Criterion**		
Model:	**PROBIT**		**PROBIT**		**PROBIT**		
Dependent Variable:	**Entry in Grid Box=1**		**Entry in Grid Box=1**		**Entry =1**		
Sample	**Sector's with Trees > 30,000 & Eligible Grid Box=1**		**Sector's with Trees > 30,000 & Eligible Grid Box=1**		**Sector's with Trees > 30,000 & Eligible Grid Box=1**		
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Spring = 1 in grid box	0.213*	0.377**			0.204*	0.374***	-3.921***
	(0.129)	(0.150)			(0.114)	(0.142)	(0.308)
Untarred Local Roads = 1 in grid box			0.378***	0.362***			
			(0.116)	(0.137)			
Roads = 1 in grid box					0.379***	0.366***	0.343**
					(0.120)	(0.137)	(0.136)
(Roads = 1 in grid box)*(Spring = 1 in grid box)							4.370***
							(0.312)
Controls							
Geographical Polynomials	no	yes	no	yes	no	yes	yes
Geographical Polynomials - Interaction	no	yes	no	yes	no	yes	yes
Pseudo R2	0.002	0.12	0.006	0.12	0.007	0.124	0.125
Number of observations	13970	13970	13970	13970	13970	13970	13970

Notes: Standard errors in ( ) are clustered at the sector level (256). ***, **, * denote statistically significance at 1, 5, 10 percent, respectively. The unit of analysis is a grid box of 1 square kilometer (Rwanda's land divided into a square kilometer grid). The grid box to be eligble for the sample need to be in a sector with at least 30,000 coffee trees and need to be on land that is not built-up, natural forest, runway or water body.

TABLE 5: COMPETITION AND EFFICIENCY - REDUCED FORM, IV AND FIRST STAGE RESULTS

Model:	**OLS**	**OLS**	**2SLS**	**OLS - FIRST STAGE**
Dependent Variable:	**Ln (Efficiency)**	**Ln (Efficiency)**	**Ln (Efficiency)**	**No. of Wet Mills within 5-10 km of Wet Mill**
Sample			**Surveyed Wet Mills only**	
	[1]	[2]	[3]	[4]
No. of Wet Mills within 5-10 km of Wet Mill	-0.009**		-0.039**	
	(0.004)		(0.018)	
Predicted Score within 5-10 km of Wet Mill		-4.168***		173.33***
		(1.390)		(22.16)
Predicted Score within 5 km of Wet Mill			3.934*	
			(2.415)	
Controls				
Entry model criteria	yes	yes	yes	yes
Geographical characteristics	yes	yes	yes	yes
Wet Mill characteristics	yes	yes	yes	yes
R2	0.1742	0.1953	0.025	0.692
Number of observations	171	171	171	171

Notes: Standard errors in ( ) are clustered at the sector level (170). ***, **, * denote statistically significance at 1, 5, 10 percent, respectively. Column 3 is 2SLS estimation with the instrument for the number of wet mills within 5-10 km of wet mill being the predicted score from the entry model within the 5-10 km radius of the wet mill conditional on controlling for the predicted score within the 5 km radius of the wet mill as well as conditioning on all geogrpahical and wet mill level level controls.

TABLE 6: COMPETITION AND OTHER OUTCOMES - REDUCED FORM AND IV RESULTS

Model:	**OLS: Reduced Form**	**2SLS**	**OLS**	**2SLS**	**OLS**	**2SLS**
Dependent Variable:	**Ln (Other Costs)**		**Ln (Prices)**		**Ln (Conversion Ratio)**	
Sample	**Surveyed Wet Mills only**		**Surveyed Wet Mills only**		**Surveyed Wet Mills only**	
	[1]	[2]	[3]	[4]	[5]	[6]
Predicted Score within 5-10 km of Wet Mill	5.509*		1.482		-0.185	
	(3.200)		(1.220)		(0.473)	
No. of Wet Mills within 5-10 km of Wet Mill		0.073**		0.002		0.003
		(0.033)		(0.009)		(0.004)
Predicted Score within 5 km of Wet Mill		-11.59**		1.985		-1.095
		(4.55)		(1.563)		(0.504)
OLS Coefficient: No. of Wet Mills within 5-10 km of Wet Mill		*0.004*		*0.008***		0.0005
		(0.011)		*(0.004)*		(0.0013)
Controls						
Entry model criteria	yes	yes	yes	yes	yes	
Geographical	yes	yes	yes	yes	yes	
Wet Mill	yes	yes	yes	yes	yes	
R2	0.1615	0.036	0.3043	0.332	0.096	0.149
Number of observations	170	170	158	158	159	159

Notes: Standard errors in ( ) are clustered at the sector level (170). ***, **, * denote statistically significance at 1, 5, 10 percent, respectively. Column 3 is 2SLS estimation with the instrument for the number of wet mills within 5-10 km of wet mill being the predicted score from the entry model within the 5-10 km radius of the wet mill conditional on controlling for the predicted score within the 5 km radius of the wet mill and all geogrpahical and wet mill level controls.