RIETI Discussion Paper Series 14-E-014

Effects of Business Networks on Firm Growth in a Cluster of Microenterprises: Evidence from rural Ethiopia

ISHIWATA Ayako
University of Tokyo

Petr MATOUS
University of Tokyo

TODO Yasuyuki
RIETI

RIETI
Research Institute of Economy, Trade & Industry, IAA

The Research Institute of Economy, Trade and Industry
http://www.rieti.go.jp/en/

RIETI Discussion Paper Series 14-E-014

March 2014

Effects of Business Networks on Firm Growth in a Cluster of Microenterprises: Evidence from rural Ethiopia*

ISHIWATA Ayako
The University of Tokyo

Petr MATOUS
The University of Tokyo

TODO Yasuyuki
RIETI and the University of Tokyo

Abstract

Poverty reduction in rural Africa necessitates diversification of income sources from agriculture to nonfarm activities. Clustering of micro-enterprises in rural areas can promote nonfarm income. This study examines the determinants of growth in sales and skill levels of microenterprises in a tailor cluster in rural Ethiopia, focusing on the role of business networks. We collected panel data, including measures of business networks through procurement, outsourcing, and financing, for three years from 136 firms, the population in the "survival" cluster. The results show that when firms are closer to the center of business networks, i.e., firms are characterized by a higher centrality measure, they are more likely to increase sales. However, although network centrality is also associated with a higher level of tailoring skills, the skill level itself has no significant effect on sales. The finding implies that consumers in the area are not concerned much about the quality of products. Therefore, while expanding business networks can promote sales and skill levels, incentives to upgrade skills are minimal.

Keywords: Survival cluster, Firm networks, Ethiopia
JEL classifications: L14, R12

RIETI Discussion Papers Series aims at widely disseminating research results in the form of professional papers, thereby stimulating lively discussion. The views expressed in the papers are solely those of the author(s), and neither represent those of the organization to which the author(s) belong(s) nor the Research Institute of Economy, Trade and Industry.

* This study is conducted as a part of the Project "Empirical Analysis on Determinants and Impacts of Formation of Firm Networks" undertaken at Research Institute of Economy, Trade and Industry (RIETI). This study utilized firm-level data collected in Ethiopia by the authors. The authors thank Masahisa Fujita, Masayuki Morikawa, Hiroyuki Nakata, Yasuyuki Sawada, and participants at the RIETI Discussion Paper seminar for helpful comments. Financial support from RIETI and JSPS Grant-in-Aid for Scientific Research is gratefully acknowledged.

1. Introduction

In less developed countries where economic activities are primarily dominated by agriculture, structural transformation from agriculture to nonfarm activities is found to reduce poverty (Christiaensen and Todo, 2013; Otsuka and Yamano, 2006; Van Den Berg and Kumbi, 2006; Barrett et al., 2001; Bryceson, 1996). This is also the case in Sub-Saharan Africa, as observed in Christiaensen, Weerdt and Todo (2013), Van Den Berg and Kumbi (2006), Barrett, Reardon and Webb (2001), and Bryceson (1996). The development of nonfarm activities can possibly be promoted by formation of industrial clusters[1] of micro-enterprises in rural areas and secondary towns, because industrial agglomeration is beneficial to regional economic activities in general (Fujita and Thisse, 2013; Marshall, 1890) as well as in less developed countries (Sonobe and Otsuka, 2011).

Marshall (1890) argues that agglomeration of firms and workers improves firm productivity and reduces production costs through availability of specialized intermediate products, spillovers of knowledge and information, and pooling of skilled labor. Henderson (2003) indeed finds significant knowledge spillover effects in the high-technology industry in the United States. Casper (2007) finds that labor mobility of senior managers in a biotechnology cluster in San Diego is associated with diffusion of their knowledge within the cluster, implying that active networks promote knowledge diffusion.

However, these studies are mostly focusing on clusters in developed countries, and studies on the impacts of networks on growth in clusters in rural areas of less developed countries, particularly those in Sub-Saharan Africa, are scarce. Industrial clusters in Sub-Saharan Africa are quite different from those in developed countries, as they are mostly stagnant (Mano et al. 2012, Yoshino 2011) and can be classified as static (Giuliani, 2005) or survival clusters (Altenburg and Meyer-Stamer, 1999). Therefore, results from developed countries cannot be applied to the survival clusters in rural Africa.

Therefore, this study examines the effect of business networks on growth of micro-enterprises in rural Africa, using firm-level panel data from a tailor cluster in a secondary town of Ethiopia. Our data are unique in the following two aspects. First, we surveyed to the population of 136 tailoring firms in the cluster so that we can mostly identify a matrix of full business collaboration networks among the

[1] Although definitions of industrial clusters can vary, this paper adopts a simple and standard definition of Schmitz and Nadvi (1999): sectoral and spatial concentrations of firms.

tailoring firms. The information of full networks is available for three years. Second, we measure the level of tailoring skills, i.e., time efficiency and product quality, utilizing a sewing test in which each tailor was asked to sew following a given line pattern. Usually, measuring skill or technology levels at the firm level is difficult, particularly for micro-enterprises in less developed countries where the amount of capital stocks is not available so that total factor productivity cannot be measured. These advantages of our dataset enable us to investigate the effect of business networks on sales and skill levels of micro-enterprises in the tailor cluster in rural Africa.

Several studies have examined the relation between networks and firm growth in Sub-Saharan Africa. For example, Chipika and Wilson (2006) reveal a positive association between networks and technological learning in four light-engineering small and medium enterprises in Zimbabwe. Akoten and Otsuka (2007) also find that family-based micro tailors in Nairobi, Kenya are more likely to grow to larger manufacturers if they have more networks with other tailors. However, these studies deal with enterprises in mega cities, rather than in rural areas, and thus are not related to transformation from agriculture to nonfarm activities in rural areas and secondary towns. Therefore, this paper is the first study that econometrically examines the effect of networks on firm growth in rural Africa, to the authors' best knowledge.

Our results show that when a given firm is connected with more firms in business collaboration networks in the cluster, it is more likely to earn larger sales and to be equipped with higher skills. However, the skill level has no significant effect on sales. These results imply that business networks contribute to sales growth directly through improving availability of inputs and sharing of works in the cluster and to skill growth possibly through knowledge diffusion. In addition, it is suggested that indifference to product quality of both consumers and producers in the localized market may reinforce each other, leading to the stagnant and survival nature of the cluster. Therefore, creating ties with outsiders, particularly traders from outside the cluster who know preferences of consumers in cities, may be helpful to develop survival clusters from the subsistence level. The implication is consistent with the findings of existing studies, such as Akoten and Otsuka (2007), Giuliani (2011, 2005), and Yamamura, Sonobe and Otsuka (2003).

The rest of the paper is organized as follows. The next section explains empirical methodologies, and Section 3 describes data used in analysis. Section 4 demonstrates and discusses empirical results. Section 4 concludes.

2. Estimation Method

Our estimation is based on a production function of enterprise *i*:

$$\ln Y_{it} = \beta_1 + \beta_2 \ln L_{it} + \beta_3 T_{it} + \beta_4 N_{it} + \beta_5 X_{it} + \beta_6 D_{it} + \varepsilon_{it}, \quad (1)$$

Y_{it}, L_{It}, and T_{it} are sales from tailoring, the number of workers, and the skill level of tailoring firm *i* in year *t*, respectively. As we will explain later, sales from tailoring can be a good proxy for value added in our study. N_i represents a measure of the centrality of business networks, as defined in detail in the next section. According to Marshall (1890) and subsequent studies, industrial agglomeration raises the production level of firms directly through improving availability of inputs and pooling of skilled labor. Because we predict that this agglomeration effect on production is larger when business networks within the agglomeration are more developed, we expect the coefficient on the network variable in the production function is positive. X_i is a vector of other control variables, including the age, years of schooling, and ethnicity of the manager of firm *i*, the firm age, and the dummy variable which takes a value of one if the manager's parent is or was working for in the garment industry (i.e., either as a tailor or a trader). To adjust year-specific shocks including price changes over time, we include in the sales function year dummies, D_{it}, when panel estimation methods are utilized. Error terms are represented by ε and v.

Usually, production functions include the amount of capital, but as is often the case in surveys to micro-enterprises in less developed countries, we do not have reliable data on total capital stocks. Although the number of sewing machines is available in the dataset, it is highly correlated with the number of workers: The correlation coefficient is 0.893, and both are one for 44 percent of firms. Therefore, to avoid multicollinearity between labor and capital, we do not include the number of sewing machines as a determinant of sales.

An obvious econometric issue in estimation of equation (1) using ordinary least squares (OLS) is possible biases due to endogeneity. For example, while expanding business networks may result in sales expansion, sales growth can also lead to a larger size of business networks. Demand for goods and labor in the local economy may have been affected by the same unobservable shock. More importantly, the skill level is possibly determined by business networks endogenously. In addition to the direct effect on production through improving availability of inputs and pooling of skilled labor argued above, business networks may indirectly increase production through improving the technology and skill level stimulated by knowledge spillovers through networks. To incorporate these endogeneity issues and the indirect effect of networks, we employ the generalized method of moments (GMM) estimation in which the number of workers, the skill level, and the network measure are assumed to be endogenous and thus are instrumented by their lagged values and the dummy variable which takes a value of one if the manager is a female. These instruments are correlated with the current endogenous variables but not with the current error term that determines production. We will later find the validity of the instruments using the Hansen *J* test. We will also check whether the instruments are not weak using the results from first-stage regressions.

To check the robustness of the results, we take the panel nature of the data into consideration and estimate equation (1) using fixed- and random-effects models. The standard Hausman test is implemented to test whether the fixed- or random-effects model is preferred.

3. Data

3.1 Description of the surveys

The data used in this study were collected by the authors in a cluster of tailoring firms in a secondary town in rural areas of Ethiopia, Asella. Asella has a population of 67,269 in 2007 and is the administrative and economic center of Arsi Zone of a population of 2,637,657 where agriculture is the considerably dominant production activity. It is located about 175 kilometers (or 110 miles) away from Addis Ababa, the capital city, and 75 kilometers (47 miles) away from Adama, the second largest city in Ethiopia. It takes more than 3 hours by bus, the only public transportation available there, from the center of Asella to the nearest large city, Adama. In particular, people in rural areas around Asella who need several hours on foot or by bus to go to the nearest town, Asella, cannot regularly go to Adama or Addis

Ababa. Therefore, Asella is quite segregated from other economies, and customers of economic activities in Asella are mostly local people including those in rural areas.

There is a market, approximately 200 meters by 300 meters wide, near the central bus station in Asella, selling many types of consumer goods from foods to farming equipment. The market is also a cluster of garment-producing firms which make, repair, or alter clothing using their sewing machines in the market area. Garment production is the major manufacturing activity is the market. In August, 2011, 2012, and 2013, we interviewed all tailoring firms observable in the market that make men's and women's clothing, such as suits, jackets, and trousers, and sell it directly to end-users.[2] A firm is defined as a business entity holding an independent business account and also autonomous in administrative planning, following the definition of Penrose (1995). There were 188 firms in 2011, 204, in 2012, and 198 in 2013, although the number varies across years because of entries and exits. This paper particularly focuses on 136 tailoring firms that existed in the market throughout the three years. The location of firms is shown in Figure 1.

From the firm surveys, we collected information on standard business information, such as sales, the number of workers, the number of sewing machines as a measure of capital stocks, and the firm's age, and characteristics of the manager of each firm, such as his/her age, years of schooling, ethnicity, religion, and gender. We also collected information on business collaboration networks by requesting to list up-to five tailoring firms with which the firm cooperates most closely within a year. Further, we asked firms about information of each of the five firms listed, such as its location (within or outside the cluster) and the type of cooperation (e.g., procurement, subcontracting, information exchange, and credit). From this information, we can fully identify business collaboration networks within the cluster. Listing firms without a complete list of firms in the cluster may lead to biases because of memory losses. However, Brewer and Webster (2000) find that people are more likely to forget those with weaker relationships and that memory losses in such surveys would not be lead to a large bias in identifying personal networks. Therefore, networks identified in this study can be considered to be those with strong ties.

In 2012 and 2013, we also conducted a sewing test to each firm, in which its major tailor is asked to sew a particular line patter given by the survey team, as shown in Figure 2, for a fixed fee of 20 birr (or

[2] We exclude from the sample other types of fabric-processing firms that produce curtains, table clothes, other clothes for interior uses, and embroideries.

approximately one US dollar). From the sewing test, we measure the skill level of tailoring firms, as we will explain in detail in the next subsection. Although we have panel data for three years, we did not conduct the sewing test in the first year, 2011, and hence can measure skill levels only for 2012 and 2013. Therefore, this study relies on the panel data for the two years.

These tailoring firms are appropriate for the study of the role of business networks in sales and skill growth, because Asella is quite segregated as explained earlier. Business networks of the tailoring firms are mostly restricted to those within the cluster, and the firms are not closely connected to firms and institutions outside the town. For example, among the 136 firms in our sample, only 5, 7, and 5 firms received support from traders outside the region (Arsi Zone) to a large, moderate, and small extent, respectively. Therefore, by collecting information on business networks within the cluster, we can cover most of their networks.

3.2 Construction of variables

Our key dependent variable is sales from tailoring. Tailoring firms mostly receive textiles from customers and make clothes, whereas 19 percent of firms sell textiles for their tailoring. However, sales from tailoring do not include earnings from such sales of textiles and hence should be roughly equal to value added. When constructing the number of workers, a key independent variable, we adjusted the number of temporary workers to the number of full-time workers using their work days in a year.

The skill level of tailoring firms is measured from the sewing test in the surveys from the two perspectives, time efficiency and product quality. First, we measured how long each tailor takes to finish the sewing. The sewing time indicates time efficiency of the tailoring firm. Second, we evaluated the quality of sewing by measuring the area surrounded by the given line pattern and the actual stitch made by the tailor (See Figure 2). Apparently, if a tailor perfectly follows the given line, the area of deviation is zero, but as the actual sewing deviates from the given line, the area becomes larger. To construct a single index of the skill level, we standardized the log of the two measures[3] so that their means are zero and their standard deviations are 1. We added the standardized sewing time and the standardized areas of deviation and further multiply by -1 so that a higher value indicates a higher skill level. Finally, we once

[3] In the case of the area of deviation, we add 0.01 before taking the log, because it is zero for some firms.

again standardize the sum of the two standardized measures to construct our index of the skill level.

Other key independent variables are related to business collaboration networks. We particularly focus on business collaboration networks through which money or information flows into a particular firm from its partner, because such inflows of money and information can be a source of sales expansion and skill improvement. Accordingly, we define a directional tie from firm *k* to firm *i* when firm *k* reported any flow of money (associated with procurement, subcontracting, or credit) or information from *k* to *i*. Then, although we restrict the number of collaborators that each firm can report up to five, the number of incoming ties to firm *i* can exceed 5. Using the term in the social network literature, this number of incoming ties for each firm is an indegree centrality measure, the simplest measure of a given firm in the network (Jackson, 2010).

In addition, to measure how close a given firm is to any other firm, we use a measure of in-closeness centrality (Freeman, 1979). More specifically, we define a normalized in-closeness centrality measure for firm *i* as follows:

$$N_i^c = \frac{\sum_{k \neq i} \left(1/d(k,i) \right)}{n-1} \times 100 \,, \tag{2}$$

where *n* is the total number of firms in the network, *d*(*k*, *i*) represents the number of collaboration links in the shortest path from firm *k* to *i*. When firm *k* or *i* is not connected with any firm, *d*(*k*,*i*) is defined to be infinity so that 1/*d*(*k*,*i*) is zero. A higher value of the in-closeness centrality measure implies that the firm can be reached from other firms with only a few links and hence is located in a more central position in the network. An advantage of the in-closeness centrality measure compared with the indegree centrality measure is that the former can capture indirect links between a given firm and other firms through directly connected firms. Therefore, whether the indegree or in-closeness centrality measure is more accurate to examine impacts of networks on production and skill levels depends on whether indirect links is of importance in addition to direct links.

3.3 Summary statistics

Summary statistics of the key variables in 2012 and 2013 are presented in Table 1. The mean and the median of the number of workers (adjusted for full-time workers) are 1.53 and 1.08, respectively. It is indeed 1 for 47 percent of firms. The mean and the median of annual sales from tailoring are 16.469 and 9,850 birr (the Ethiopian currency unit), or 902 and 540 US dollars, using the average exchange rate in 2012 and 2013. Therefore, the tailoring firms in our dataset are mostly micro-enterprises with earnings at the subsistence level. In addition, the total employment of the 136 firms in 2011, 2012, and 2013 was 191, 215, and 220, respectively, and therefore, this tailor cluster can be considered as stagnant and barely surviving.

The firms are mostly old: The average firm age is 17.23 years and its median is 13. This may reflect the fact that many firms have been family-operated and succeeded from a generation to another. For 23 percent of firms, a parent of the manager was in the garment industry, i.e., either a tailor or a garment trader. The manager of the average firm is 41 years and has 9 years of schooling. 38 percent of managers are Amhara, 39 percent Oromo, and others are in other ethnic groups such as Grage.

The standardized skill level has a mean of zero and a standard deviation of one, as defined. Figure 3 shows the distribution of the two components of the skill level, sewing time to measure time efficiency, and areas of deviation from the given line to measure product quality. For both measures, there are some firms with exceptionally bad skills. The two measures are not closely related with each other, implying that some firms take time to improve the quality.

Looking at the network variables, we find that the average indegree centrality measure is 1.56: i.e., firms are listed by 1.77 firms as a collaborator on average. Figure 4 indicates that the two measures of business networks, the indegree and in-closeness centrality measures, are positively correlated. The distribution is quite skewed toward zero. 32 percent of firm-year observations are not listed by any firm, and 27 percent are named by only one firm. There is one exceptional firm which an indegree of 10, while this firm’s number of workers is just 1.08. The full network in the cluster in 2013 is shown in Figure 5.

4. Estimation Results

4.1 Benchmark results

The results from our benchmark regressions are shown in Table 2. We started with OLS estimations in columns (1) and (2) using the indegree and in-closeness centrality measures in separate regressions to avoid multicollinearity between the two. In the OLS estimations, we use data for 2012 and 2013, and thus the number of observations is 272 (= 136*2). In columns (3) and (4), we implemented the GMM estimations explained in Section 2. In the GMM estimations, because the lagged endogenous variables are used as instruments, the number of observations is 136. The p values of the Hansen J statistics shown in the bottom row indicate that we cannot reject the hypothesis that the error term and the instruments are orthogonal. In addition, Table 3 shows results from first-stage estimations of the GMM using OLS to find correlation between the endogenous variables and the instruments. Columns (1)-(3) of Table 3 correspond to column (3) of Table 2, whereas columns (4)-(6) of Table 3 to column (4) of Table 2. In each regression, at least one of the excluded instruments has a significant effect on the dependent variable, one of the endogenous variables. Also, the F statistics from the regressions are either above 10 or very close to 10, the threshold value suggested by Stock and Yogo (2002). Therefore, we conclude that in both GMM estimations, the instruments are valid and not weak.

The GMM results find that the number of workers has a statistically significant effect on sales and that the size of the coefficients, 0.803 and 0.895, that indicate the elasticity of labor in the production function is reasonable. However, the effect of the skill level is found to be insignificant in both GMM estimations. We will discuss the implication of this result in detail later. The effect of the indegree centrality measure on sales is significant at the 5-percent level, while the effect of the in-closeness centrality measure is insignificant. The contrasting result is obtained probably because in-closeness centrality covers indirect links between firms whereas indegree centrality only measures direct links. Therefore, the results imply that sales are affected by direct collaboration links with other firms, rather than indirect links within the whole cluster.

Among other control variables, the age of the manager of each firm has a negative and significant effect, implying that younger managers are likely to earn more. The effect of managers' age is large in magnitude, as the GMM results imply that when the age of the manager doubles, sales decrease by more than 70 percent. When the parent of the manager has any experience in the garment industry, sales increase by 33 percent. The dummy of Amhara, a major ethnic group, has a large positive effect. However, because we control for network centrality in this regression and find no significant effect on the network measures in the first-stage regression (Table 3), this positive effect of a particular ethnic group may not reflect an effect of business networks based on ethnicity.

As we have seen, Table 3 shows correlation between the endogenous variables and the instruments in the GMM estimations. Column (2) indicates that the indegree centrality measure has a positive effect on the skill level. However, the in-closeness centrality measure is not correlated with the skill level (column [5]), probably because indirect collaboration links are not helpful to skill improvement.

These results imply that expansion of business networks contribute to sales growth directly through improving availability of inputs and sharing of works in the cluster and to skill growth possibly through knowledge diffusion. However, because the skill level has no significant effect on sales, as column (3) of Table 2 shows, the skill improvement through business networks does not actually contribute to sales growth in this cluster.

4.2 Robustness checks

To check the robustness of the benchmark results, we run fixed- and random-effects estimations. Using the standard Hausman test, we test the hypothesis that the error term is uncorrelated with the independent variables so that the assumption of the random-effects model is satisfied. The *p* values of the Hausman statistics shown in the bottom row indicate that we cannot reject the hypothesis. Therefore, throughout the panel estimations, we rely on the results from the random-effect estimation that yield smaller standard errors. Columns (3) and (4) of Table 4 show that both indegree and in-closeness centrality measures have a positive and significant effect on sales, while the skill level has no significant effect. Moreover, the two centrality measures positively affect the skill level, according to columns (7) and (8). These results from the panel estimations are mostly consistent with those from the GMM, except for the positive effect of the in-closeness centrality measure on sales and skills.

In addition, we experiment with three alternative sets of estimations. First, we assume that the number of workers is exogenous because it may be fixed for a short period of time for the mostly family-operated micro-enterprises. Second, we use the second lag of the number of workers and the networks variables as instruments, utilizing data in 2011. Third, we drop from the sample outliers whose sales per worker or the skill level is smaller than its mean minus three times its standard deviation or larger than its mean plus three times its standard deviation. These alternative estimations yield the same conclusions, although the results are not shown here for brevity but available upon request from the authors.

4.3 Disaggregating skills

As described in Section 3.2, the measure of the skill level used so far contains two aspects of skills, time efficiency measured by the time of sewing the given pattern shown in Figure 2 and product quality measured by the area of deviation from the given pattern. Now, we distinguish the two and run separate sets of estimations, as in Tables 2 and 3. According to the results using time efficiency as a skill level shown in Table 5, the indegree centrality measure has a positive and significant effect on sales and time efficiency, and time efficiency has no significant effect on sales (columns [1] and [3]). However, the in-closeness centrality measure has no significant effect on either sales or time efficiency (columns [5] and [7]). These results are fully consistent with the benchmark results. In Table 6 where product quality is used as a skill measure, we find that both indegree and in-closeness measures positively affect sales, while either does not affect product quality and that product quality has no significant effect on sales. These results are not fully coincide with the benchmark results where the integrated skill level is used, but the main results that business networks positively affect sales while skill levels do not necessarily improve sales still hold.

5. Discussion and Conclusion

This paper examines how business networks in a survival cluster of microenterprises in rural Ethiopia affect their sales and skill levels, using a firm-level panel data for the population of 136 tailoring firms. We collected unique data, including business collaboration networks among firms in the cluster and skill levels measured from a sewing test. We employ GMM estimations to correct for endogeneity biases using lagged variables and other exogenous variables as instruments. Our results show that when a given firm is connected with more firms in business collaboration networks in the cluster, or when it has a higher degree centrality measure, it is more likely to earn larger sales and to be equipped with higher skills. However, the skill level has no significant effect on sales. These results imply that business networks contribute to sales growth directly through improving availability of inputs and sharing of works in the cluster and to skill growth possibly through knowledge diffusion. These results mostly hold when we employ different estimation methods or measures of the skill level.

Our finding that the skill level does not improve sales is not standard and needs explanation. We interpret this finding as showing that because customers of the clustering firms, mostly local consumers in the rural region of Ethiopia, are not seriously concerned about the quality of products, firms are also less concerned about quality. In the sewing test conducted during the survey, 36 percent of firms made tangles, thread cut, or textile tension that are less likely to be accepted in markets in cities. Business networks contribute to time efficiency in tailoring but not to product quality in particular. As a result, firms in improvement in product quality in the cluster is minimal, and hence firms remain at the subsistence level. The indifference to product quality of both consumers and producers in the localized market may reinforce each other, leading to the stagnant and survival nature of the cluster.

One possible way to overcome this stagnant state is to connect micro-enterprises in the rural cluster with urban markets where consumers are more concerned about product quality. Once producers realize the importance of product quality, they may improve their product quality by fully utilizing their business networks to develop from the subsistence level to a higher stage. This implication is consistent with several existing studies. For example, Akoten and Otsuka (2007), Giuliani (2011, 2005), and Yamamura, Sonobe and Otsuka (2003) respectively point to the importance of traders and technological gatekeepers who bring information on technology and consumers' preferences from outside to the cluster in cluster development, using case studies of a tailor cluster in Nairobi, a wine cluster in Chile, and a garment cluster in the post-WWII Japan. In other words, strengthening business networks within the cluster may be sufficient for the cluster to develop beyond the subsistence level.

Reference

Akoten J. E. and Otsuka K. (2007). From Tailors to Mini-Manufacturers: The Role of Traders in the Performance of Garment Enterprises in Kenya. *Journal of African Economies.* 16 (4), 564-95.

Altenburg T. and Meyer-Stamer J. (1999). How to Promote Clusters: Policy Experiences from Latin America. *World Development.* 27 (9), 1693-713.

Barrett C. B., Reardon T. and Webb P. (2001). Nonfarm Income Diversification and Household Livelihood Strategies in Rural Africa: Concepts, Dynamics, and Policy Implications. *Food policy.* 26 (4), 315-31.

Brewer D. D. and Webster C. M. (2000). Forgetting of Friends and Its Effects on Measuring Friendship Networks. *Social Networks.* 21 (4), 361-73.

Bryceson D. F. (1996). Deagrarianization and Rural Employment in Sub-Saharan Africa: A Sectoral Perspective. *World Development.* 24 (1), 97-111.

Casper S. (2007). How Do Technology Clusters Emerge and Become Sustainable?: Social Network Formation and Inter-Firm Mobility within the San Diego Biotechnology Cluster. *Research Policy.* 36 (4), 438-55.

Chipika S. and Wilson G. (2006). Enabling Technological Learning among Light Engineering Smes in Zimbabwe through Networking. *Technovation.* 26 (8), 969-79.

Christiaensen L. and Todo Y. (2013). Poverty Reduction During the Rural–Urban Transformation–the Role of the Missing Middle. *World Development.*).

Christiaensen L., Weerdt J. and Todo Y. (2013). Urbanization and Poverty Reduction: The Role of Rural Diversification and Secondary Towns. *Agricultural Economics.* 44 (4-5), 435-47.

Freeman L. C. (1979). Centrality in Social Networks Conceptual Clarification. *Social Networks.* 1 (3), 215-39.

Fujita M. and Thisse J.-F. (2013). *Economics of Agglomeration: Cities, Industrial Location, and Globalization.* Cambridge University Press.

Giuliani E. (2005). Cluster Absorptive Capacity Why Do Some Clusters Forge Ahead and Others Lag Behind? *European Urban and Regional Studies.* 12 (3), 269-88.

Giuliani E. (2011). Role of Technological Gatekeepers in the Growth of Industrial Clusters: Evidence from Chile. *Regional Studies.* 45 (10), 1329-48.

Henderson J. V. (2003). Marshall's Scale Economies. *Journal of urban economics.* 53 (1), 1-28.

Jackson M. O. (2010). *Social and Economic Networks.* Princeton University Press.

Mano Y., Iddrisu A., Yoshino Y. and Sonobe T. (2012). How Can Micro and Small Enterprises in Sub-Saharan Africa Become More Productive? The Impacts of Experimental Basic Managerial Training. *World Development.* 40 (3), 458-68.

Marshall A. (1890). *The Principles of Economics.* London: Macmillan.

Penrose E. T. (1995). *The Theory of the Growth of the Firm.* Oxford University Press.

Schmitz H. and Nadvi K. (1999). Clustering and Industrialization: Introduction. *World Development.* 27 (9), 1503-14.

Sonobe T. and Otsuka K. (2011). *Cluster-Based Industrial Development: A Comparative Study of Asia and Africa.* Palgrave Macmillan.

Stock J. H. and Yogo M. (2002). Testing for Weak Instruments in Linear Iv Regression *NBER Working Paper.* 284).

Van Den Berg M. and Kumbi G. E. (2006). Poverty and the Rural Nonfarm Economy in Oromia, Ethiopia. *Agricultural Economics.* 35 (s3), 469-75.

Yamamura E., Sonobe T. and Otsuka K. (2003). Human Capital, Cluster Formation, and International Relocation: The Case of the Garment Industry in Japan, 1968–98. *Journal of Economic Geography.* 3 (1), 37-56.

Yoshino Y. (2011). *Industrial Clusters and Micro and Small Enterprises in Africa: From Survival to Growth.* Washington D.C.: World Bank Publications.

Figure 1: Map of the Cluster

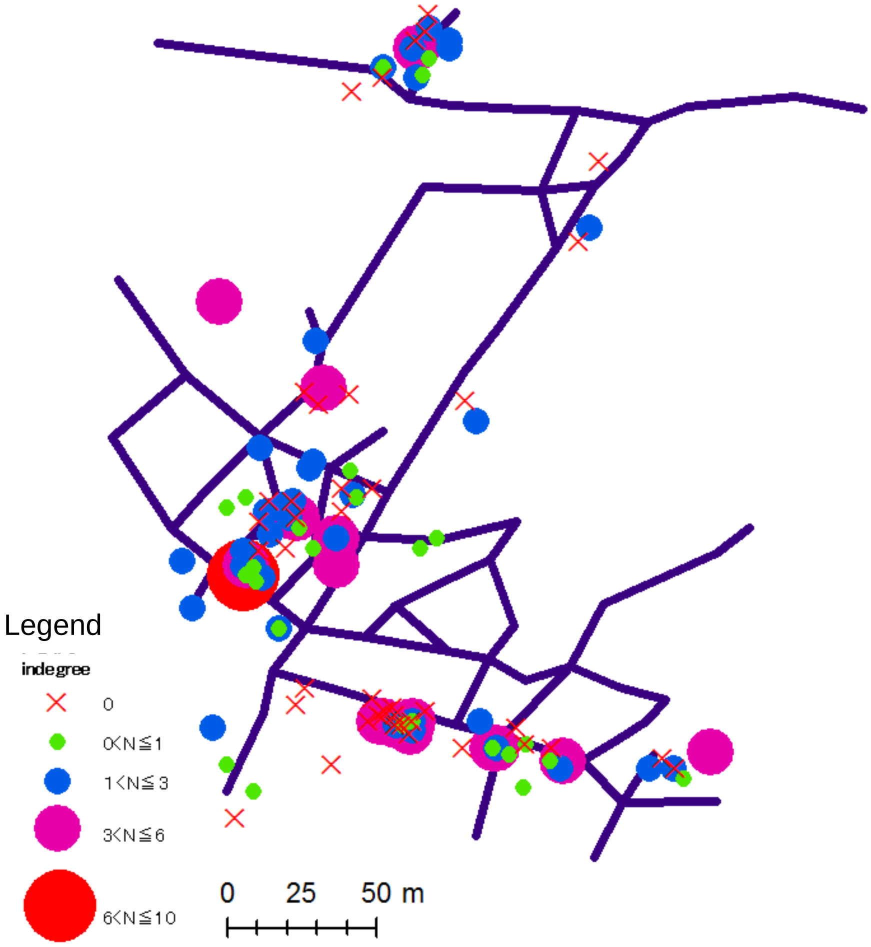


Figure 2: Sewing Test

(1) An example of good sewing (2) An example of bad sewing

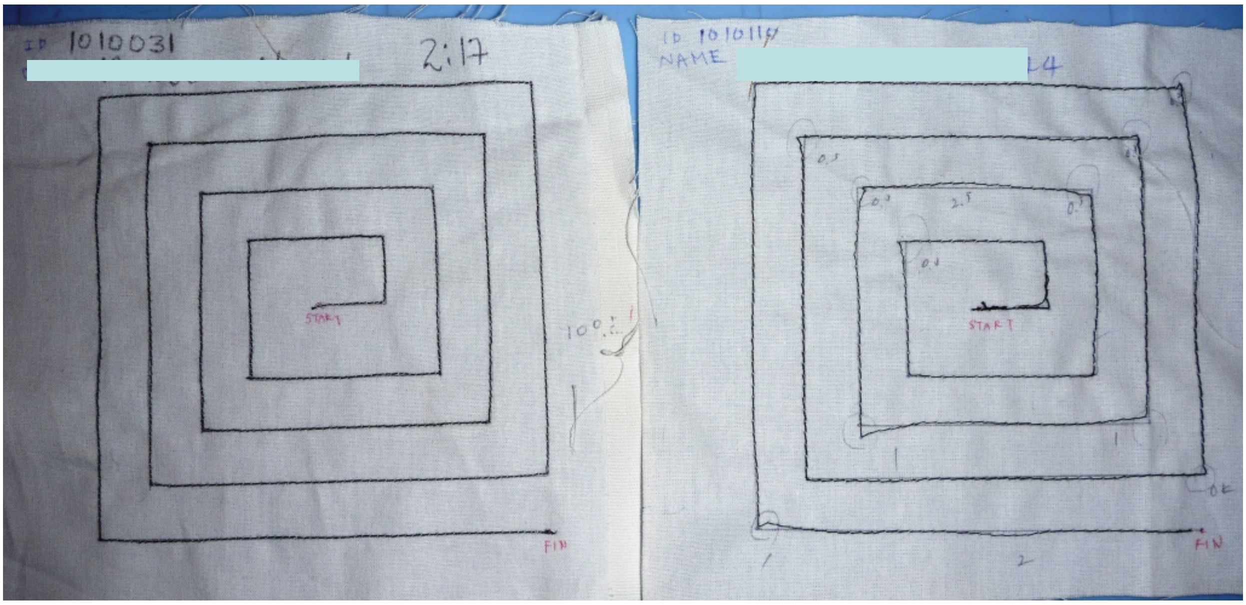


Figure 3: Distribution of Skill Levels

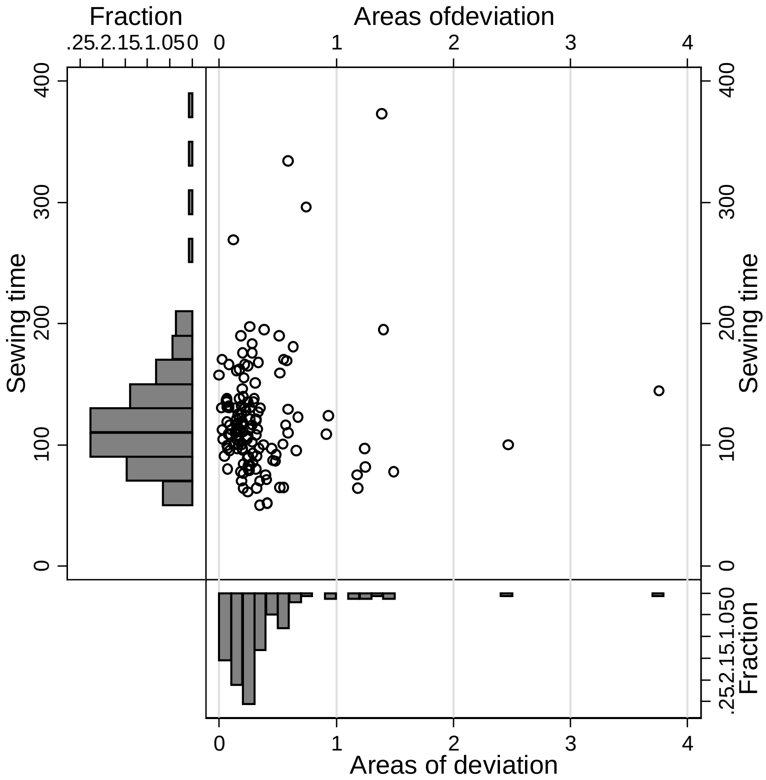


Figure 4: Distribution of Network Measures

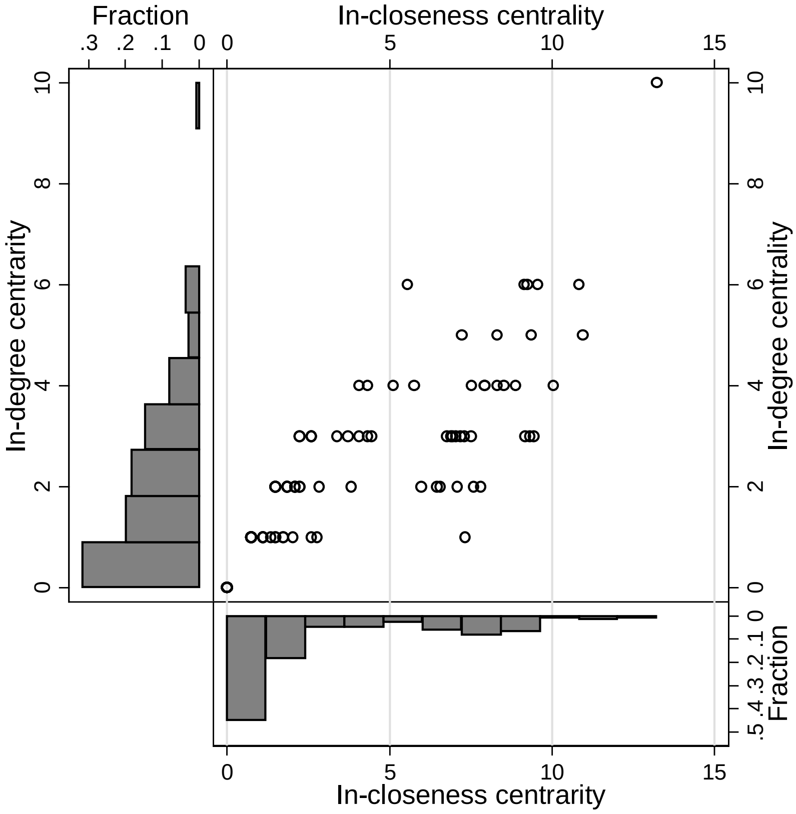
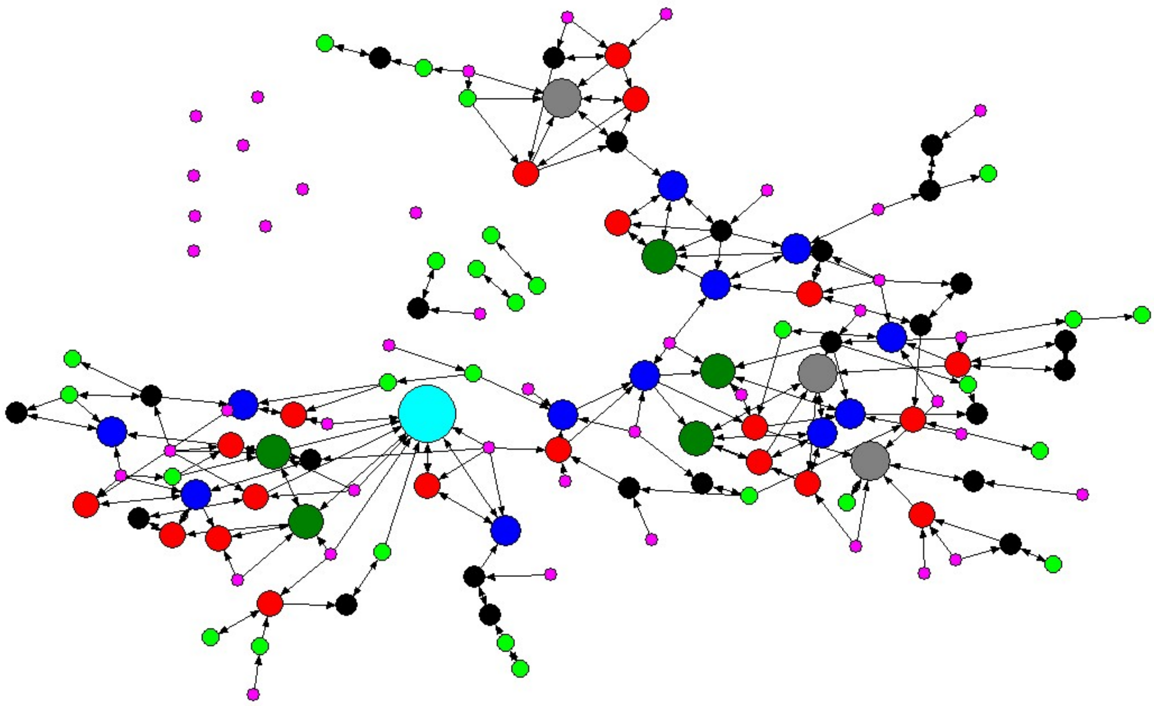


Figure 5: Full Business Networks in the Cluster



Note: The size of each circle depicts the indegrees of a given firm, i.e., the number of firms which reported that they had a collaboration relation with the firm.

Table 1: Summary Statistics

	Mean	S.D.	Median	Min.	Max.
Sales	16,469	36,222	9,850	396	648,000
-- in logs	9.10	1.07	9.20	5.98	13.38
Number of workers	1.53	1.34	1.08	1	13
-- in logs	0.28	0.45	0.08	0	2.56
Standardized skill level	0.00	1.00	0.09	-3.91	2.03
Sewing time (seconds)	127.19	47.32	119	50	373
Areas of deviation (cm^2)	0.40	0.55	0.23	0	3.79
Indegree centrality	1.56	1.75	1.00	0	11
-- in logs	0.74	0.63	0.69	0	2.48
In-closeness centrality	3.20	4.20	1.48	0	22.17
-- in logs	-0.83	2.85	0.40	-4.61	3.10
Firm age	17.23	13.37	13.00	1	53
-- in logs	2.44	1.00	2.56	0	3.97
Manager’s characteristics					
Age	41.27	13.64	40.00	18	78
-- in logs	3.66	0.34	3.69	2.89	4.36
Years of schooling	9.06	3.11	10.00	0	13
-- in logs	2.21	0.55	2.40	0	2.64
Dummy for parent in garment industry	0.23	0.42	0.00	0	1
Amhara	0.38	0.49	0.00	0	1
Oromo	0.39	0.49	0.00	0	1
Female	0.08	0.28	0.00	0	1

Note: These are the summary statistics for the 272 firm-year observations for years 2012 and 2013.

Table 2: Benchmark Results

Dependent variable: Sales in logs

	(1)	(2)	(3)	(4)
	OLS	OLS	GMM	GMM
Number of workers	0.823***	0.869***	0.803***	0.895***
	(0.105)	(0.105)	(0.204)	(0.203)
Skill level	0.0491	0.0591	0.0941	0.105
	(0.0515)	(0.0524)	(0.128)	(0.130)
Indegree centrality	0.238***		0.344**	
	(0.0763)		(0.167)	
In-closeness centrality		0.0369**		0.0635
		(0.0166)		(0.0450)
Firm age	0.0993	0.0995	-0.0107	-0.0124
	(0.0842)	(0.0844)	(0.125)	(0.122)
Age	-0.940***	-0.921***	-0.758**	-0.738**
	(0.248)	(0.249)	(0.377)	(0.373)
Years of schooling	-0.129	-0.132	-0.157	-0.174
	(0.0900)	(0.0907)	(0.121)	(0.118)
Parent in garment industry	0.430***	0.425***	0.337**	0.332**
	(0.113)	(0.113)	(0.155)	(0.153)
Amhara	0.364***	0.362***	0.440***	0.438***
	(0.113)	(0.114)	(0.152)	(0.150)
Oromo	0.147	0.150	0.182	0.190
	(0.114)	(0.116)	(0.162)	(0.161)
Observations	272	272	136	136
R-squared	0.431	0.421	0.439	0.439
Hansen *J* statistic (*p* value)			0.192	0.233

Notes: Standard errors are in parentheses. *, **, and *** signify statistical significance at the 10-, 5-, and 1- percent level. The number of workers, the indegree and in-closeness centrality measures, and the firm age are in logs. The age and years of schooling are those of the manager of each firm and in logs. "Parent in garment industry" is the dummy which takes a value of one if the parent of the manager was working for the garment industry either as a tailor or a trader. Amhara and Oromo are the dummies for the two major ethnic groups.

Table 3: Correlation between Endogenous Variables and Instruments

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	Number of workers	Skill level	Indegree centrality	Number of workers	Skill level	In-closeness centrality
Number of workers (lag)	0.854***	0.0506	0.287***	0.850***	0.0685	0.947***
	(0.0768)	(0.156)	(0.0857)	(0.0788)	(0.153)	(0.339)
Skill level (lag)	0.0165	0.531***	0.0202	0.0137	0.529***	0.124
	(0.0287)	(0.0908)	(0.0400)	(0.0289)	(0.0893)	(0.206)
Indegree centrality (lag)	0.0590	0.185*	0.723***			
	(0.0419)	(0.109)	(0.0552)			
In-closeness centrality (lag)				0.0169**	0.0365	0.590***
				(0.00746)	(0.0248)	(0.0680)
Female	0.132	-0.247	0.171	0.126	-0.272	0.521
	(0.0909)	(0.320)	(0.157)	(0.0932)	(0.315)	(0.959)
Age	0.271*	-0.820**	-0.0950	0.271*	-0.808**	-0.675
	(0.138)	(0.330)	(0.210)	(0.138)	(0.330)	(1.035)
Years of schooling	0.0832**	0.102	-0.0466	0.0799**	0.100	-0.0941
	(0.0337)	(0.171)	(0.0584)	(0.0327)	(0.170)	(0.283)
Firm age	-0.0356	0.149	-0.0492	-0.0406	0.143	-0.208
	(0.0475)	(0.113)	(0.0695)	(0.0467)	(0.113)	(0.339)
Parent in garment industry	0.111	-0.174	-0.0328	0.114	-0.174	-0.417
	(0.0701)	(0.174)	(0.0882)	(0.0699)	(0.173)	(0.434)
Amhara	-0.000408	-0.0418	0.0206	-0.00347	-0.0462	0.0278
	(0.0645)	(0.166)	(0.0986)	(0.0643)	(0.167)	(0.502)
Oromo	-0.0363	0.0254	-0.0264	-0.0421	0.0195	-0.257
	(0.0596)	(0.180)	(0.0922)	(0.0602)	(0.178)	(0.449)
Observations	136	136	136	136	136	136
R-squared	0.661	0.472	0.633	0.665	0.469	0.477
F statistic	14.81	9.706	42.29	14.82	9.799	20.36

Notes: Standard errors are in parentheses. *, **, and *** signify statistical significance at the 10-, 5-, and 1- percent level. The number of workers, the indegree and in-closeness centrality measures, and the firm age are in logs. The age and years of schooling are those of the manager of each firm and in logs. “Parent in garment industry” is the dummy which takes a value of one if the parent of the manager was working for the garment industry either as a tailor or a trader. Amhara and Oromo are the dummies for the two major ethnic groups.

Table 4: Robustness Checks

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable	Sales in logs				Skill level			
	Fixed effects	Fixed effects	Random effects	Random effects	Fixed effects	Fixed effects	Random effects	Random effects
Number of workers	0.347**	0.337**	0.652***	0.676***				
	(0.163)	(0.163)	(0.111)	(0.111)				
Skill level	0.0242	0.0270	0.0370	0.0445				
	(0.0564)	(0.0564)	(0.0459)	(0.0462)				
Indegree centrality	0.0716		0.218***					
	(0.110)		(0.0750)					
In-closeness centrality		0.00165		0.0300*				
		(0.0209)		(0.0157)				
Age	1.167	1.813	-1.077***	-1.076***	2.373	2.242	-1.386***	-1.386***
	(6.769)	(6.754)	(0.328)	(0.332)	(10.37)	(10.39)	(0.338)	(0.335)
Years of schooling	-2.411	-2.928	-0.116	-0.116	1.742	1.604	0.114	0.109
	(4.974)	(4.981)	(0.130)	(0.132)	(7.732)	(7.651)	(0.134)	(0.133)
Firm age	0.447	0.462	0.146	0.152	-0.466	-0.476	0.181	0.180
	(0.454)	(0.457)	(0.106)	(0.107)	(0.707)	(0.708)	(0.112)	(0.111)
Parent in garment industry			0.428***	0.421***			0.111	0.114
			(0.158)	(0.160)			(0.162)	(0.161)
Amhara			0.365**	0.365**			0.107	0.108
			(0.166)	(0.168)			(0.174)	(0.172)
Oromo			0.145	0.149			0.0169	0.0197
			(0.165)	(0.167)			(0.171)	(0.169)
Number of workers (lag)					0.0612	0.0667	0.207	0.200
					(0.246)	(0.248)	(0.139)	(0.139)
Skill level (lag)								
Indegree centrality (lag)					0.0116		0.190**	
					(0.132)		(0.0912)	
In-closeness centrality (lag)						-0.00201		0.0429**
						(0.0259)		(0.0189)
Female							-1.177***	-1.190***
							(0.252)	(0.249)
Observations	272	272	272	272	272	272	272	272
R-squared	0.306	0.304	0.282	0.274	0.005	0.005	0.000	0.000
Hausman statistic (*p* value)			0.188	0.207			0.402	0.355

Notes: Standard errors are in parentheses. *, **, and *** signify statistical significance at the 10-, 5-, and 1- percent level. The number of workers, the indegree and in-closeness centrality measures, and the firm age are in logs. The age and years of schooling are those of the manager of each firm and in logs. “Parent in garment industry” is the dummy which takes a value of one if the parent of the manager was working for the garment industry either as a tailor or a trader. Amhara and Oromo are the dummies for the two major ethnic groups.

Table 5: Time Efficiency

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	GMM		First stage		GMM		First stage	
Dependent variable	Sales	Number of workers	Time efficiency	Indegree centrality	Sales	Number of workers	Time efficiency	In-closeness centrality
Number of workers	0.838***				0.952***			
	(0.199)				(0.193)			
Skill level	0.114				0.122			
	(0.112)				(0.111)			
Indegree centrality	0.303*							
	(0.166)							
In-closeness centrality					0.0447			
					(0.0443)			
Age	-0.856**	0.267**	-0.298	-0.143	-0.860**	0.267**	-0.302	-1.050
	(0.364)	(0.133)	(0.348)	(0.207)	(0.355)	(0.134)	(0.347)	(1.006)
Years of schooling	-0.181	0.0846**	0.280*	-0.0453	-0.198	0.0811**	0.263*	-0.0907
	(0.126)	(0.0337)	(0.144)	(0.0579)	(0.123)	(0.0328)	(0.138)	(0.282)
Firm age	0.00194	-0.0357	0.0854	-0.0426	0.00302	-0.0404	0.0624	-0.161
	(0.127)	(0.0478)	(0.125)	(0.0690)	(0.122)	(0.0467)	(0.124)	(0.339)
Parent in garment industry	0.330**	0.114	-0.0246	-0.0238	0.315**	0.116	-0.0126	-0.348
	(0.151)	(0.0701)	(0.158)	(0.0861)	(0.147)	(0.0700)	(0.156)	(0.421)
Amhara	0.408***	0.00124	0.169	0.0244	0.404***	-0.00201	0.153	0.0508
	(0.154)	(0.0652)	(0.169)	(0.0981)	(0.152)	(0.0650)	(0.169)	(0.501)
Oromo	0.160	-0.0357	0.262	-0.0278	0.169	-0.0416	0.233	-0.275
	(0.159)	(0.0594)	(0.171)	(0.0930)	(0.157)	(0.0601)	(0.167)	(0.456)
Number of workers (lag)		0.856***	-0.139	0.297***		0.852***	-0.155	1.009***
		(0.0770)	(0.185)	(0.0884)		(0.0790)	(0.181)	(0.347)
Skill level (lag)		0.0158	0.544***	-0.0193		0.0118	0.524***	-0.191
		(0.0224)	(0.0780)	(0.0436)		(0.0225)	(0.0744)	(0.237)
Indegree centrality (lag)		0.0595	0.298**	0.728***				
		(0.0415)	(0.121)	(0.0541)				
In-closeness centrality (lag)						0.0169**	0.0838***	0.605***
						(0.00742)	(0.0264)	(0.0671)
Female		0.129	-0.292	0.123		0.121	-0.328	0.147
		(0.0840)	(0.326)	(0.158)		(0.0856)	(0.326)	(0.955)
Observations	136	136	136	136	136	136	136	136
R-squared	0.441	0.661	0.445	0.633	0.445	0.665	0.464	0.480
Hansen *J*/*F* statistic		14.87	9.466	39.18		14.89	9.512	18.78

Notes: Standard errors are in parentheses. *, **, and *** signify statistical significance at the 10-, 5-, and 1- percent level. The number of workers, the indegree and in-closeness centrality measures, and the firm age are in logs. The age and years of schooling are those of the manager of each firm and in logs. “Parent in garment industry” is the dummy which takes a value of one if the parent of the manager was working for the garment industry either as a tailor or a trader. Amhara and Oromo are the dummies for the two major ethnic groups.

Table 6: Product Quality

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	GMM		First stage		GMM		First stage	
Dependent variable	Sales	Number of workers	Product quality	Indegree centrality	Sales	Number of workers	Product quality	In-closeness centrality
Number of workers	0.801***				0.896***			
	(0.214)				(0.215)			
Skill level	0.0105				0.0128			
	(0.167)				(0.161)			
Indegree centrality	0.388**							
	(0.158)							
In-closeness centrality					0.0766*			
					(0.0416)			
Age	-0.930**	0.252*	-0.872***	-0.0897	-0.914**	0.257**	-0.868***	-0.594
	(0.378)	(0.128)	(0.322)	(0.215)	(0.373)	(0.128)	(0.317)	(1.075)
Years of schooling	-0.149	0.0840**	-0.126	-0.0502	-0.166	0.0803**	-0.111	-0.126
	(0.120)	(0.0334)	(0.160)	(0.0572)	(0.117)	(0.0324)	(0.159)	(0.269)
Firm age	0.0189	-0.0332	0.120	-0.0484	0.0190	-0.0388	0.139	-0.209
	(0.117)	(0.0467)	(0.103)	(0.0715)	(0.115)	(0.0458)	(0.104)	(0.355)
Parent in garment industry	0.334**	0.115	-0.191	-0.0386	0.336**	0.117*	-0.202	-0.470
	(0.156)	(0.0697)	(0.206)	(0.0879)	(0.156)	(0.0695)	(0.210)	(0.431)
Amhara	0.415***	0.00137	-0.212	0.0169	0.412***	-0.00232	-0.202	-0.00692
	(0.155)	(0.0659)	(0.197)	(0.0986)	(0.153)	(0.0655)	(0.194)	(0.504)
Oromo	0.184	-0.0367	-0.221	-0.0281	0.192	-0.0426	-0.201	-0.272
	(0.165)	(0.0593)	(0.205)	(0.0928)	(0.163)	(0.0598)	(0.203)	(0.451)
Number of workers (lag)		0.858***	0.226	0.283***		0.852***	0.273	0.898***
		(0.0771)	(0.276)	(0.0851)		(0.0791)	(0.271)	(0.332)
Skill level (lag)		0.00364	0.407***	0.0414		0.00434	0.408***	0.320*
		(0.0266)	(0.0861)	(0.0372)		(0.0265)	(0.0852)	(0.190)
Indegree centrality (lag)		0.0611	-0.0373	0.722***				
		(0.0410)	(0.133)	(0.0542)				
In-closeness centrality (lag)						0.0174**	-0.0332	0.592***
						(0.00716)	(0.0308)	(0.0655)
Female		0.114	-0.0873	0.176		0.112	-0.105	0.608
		(0.0852)	(0.354)	(0.153)		(0.0856)	(0.365)	(0.951)
Observations	136	136	136	136	136	136	136	136
R-squared	0.436	0.660	0.270	0.636	0.434	0.664	0.278	0.487
Hansen *J*/*F* statistic		14.91	3.770	40.66		14.86	3.870	24.48

Notes: Standard errors are in parentheses. *, **, and *** signify statistical significance at the 10-, 5-, and 1- percent level. The number of workers, the indegree and in-closeness centrality measures, and the firm age are in logs. The age and years of schooling are those of the manager of each firm and in logs. "Parent in garment industry" is the dummy which takes a value of one if the parent of the manager was working for the garment industry either as a tailor or a trader. Amhara and Oromo are the dummies for the two major ethnic groups.