How do output prices and physical productivity respond to enterprise clustering? Evidence from a developing country*

Eyerusalem Siba† Måns Söderbom‡ Arne Bigsten§ Mulu Gebreeyesus¶

February 2013

Abstract

We use census panel data on Ethiopian manufacturing firms to investigate how enterprise clustering in local markets affects firm-level output prices and physical productivity. We find a negative and statistically significant relationship between the density of firms that produce a given product in a given location and the local price of that product. We further find a positive and statistically significant relationship between the density of firms that produce a given product in a location and the physical productivity of same-product firms in the location. These results are consistent with the notion that increased clustering of firms generates higher competitive pressure and positive externalities. The net effect on firm-level revenues is close to zero, suggesting that firms have weak incentives to agglomerate endogenously. Across firms that produce different products, we find no statistically significant relationship between enterprise clustering and firm-level output prices and productivity. We also find no clustering effects across towns. Our results thus suggest that while clustering can impact firm performance, the advantages are narrow in scope.

*This paper is one of a series of studies of industrial development in Africa to be published jointly by the Africa Growth Initiative at Brookings and UNU-WIDER under their joint project Learning to Compete (L2C). We would like to thank Jean-Luis Arcand, Markus Eberhardt, Marcel Fafchamps, John Haltiwanger, Richard Kneller, David McKenzie, John Page, Sandra Poncet, Francis Teal, Katrin Tinn, seminar participants in Cleremont-Ferrand, Copenhagen, Gothenburg, Navarra, Nottingham, Oxford, Paris, Trondheim, and participants in the CSAE Conference on Economic Development in Africa in Oxford 2011, and the Nordic Conference on Development Economics in Copenhagen 2011 for very helpful comments. All errors are our own.

†University of Gothenburg, Sweden. E-mail: Eyerusalem.siba@economics.gu.se.

‡University of Gothenburg, Sweden. Corresponding author. Telephone number: +46 31 7864332. Fax number: +46 31 7861326. E-mail: mans.soderbom@economics.gu.se.

§University of Gothenburg, Sweden. E-mail: arne.bigsten@economics.gu.se.

¶UNU-MERIT, Maastricht, the Netherlands. E-mail: gebreeyesus@merit.unu.edu.

1. Introduction

Starting with Marshall (1920), many economists have argued that geographical clustering of enterprises can be a source of improved firm performance.[1] Several mechanisms have been highlighted in the literature, for example information spillovers, better access to skilled labor, lower transaction costs, and higher competitive pressure. Numerous studies provide empirical evidence indicating that increased clustering of firms has been an important driver of growth in the USA and Europe (e.g. Glaeser et al., 1992, Henderson et al., 1995, Henderson, 1997, Combes, 2000, De Lucio et al., 2002; Blien et al., 2006). For low income countries, however, the effects of firm clustering on firm performance have not been as thoroughly documented, and for the world's poorest region, Sub-Saharan Africa, what little evidence exists comes in the form of case studies or studies based on cross-sectional data.[2]

In this paper we use census panel data on formal Ethiopian manufacturing enterprises to investigate how changes in the density of firms in a local market impact local firm-level output prices and physical productivity, controlling for a wide range of unobservable factors. Ethiopia's manufacturing sector provides an interesting setting in which to study these mechanisms. The country is large and populous, transport costs are high, and problems posed by imperfect information are common. In such an environment, the competitive threat posed by firms based outside the own location is low, while at the same time the diffusion of new ideas, practices, and technologies across locations is likely limited and slow. Within local markets, however, changes in the density of firms may have a strong impact on firm performance, if effects of competition, knowledge spillovers and other externalities are important.

That our data enable us to analyze separately the effects on firm-level output prices and firm-level physical productivity is a significant advantage. There is a growing recognition in the literature on firm

[1] Sonobe and Otsuka (2006, p.4) define a cluster as "the geographical concentration or localization of enterprises producing similar or closely related goods in a small area". Porter (1990, p. 18) defines it as a "geographical concentration of interconnected companies and institutions in a particular field". Swann et al (1998, p 1) define it as "a large group of firms in related industries at a particular location". Schmitz and Nadvi (1999) simply define industrial cluster as "sectoral and spatial concentration of firms."

[2] See Sonobe and Otsuka (2011) for a case study of cluster-based industrial development in Africa and Asia, and Fafchamps and Söderbom (2013) for a descriptive study of the role of business networks for diffusion of new technology and business practices in Ethiopia and the Sudan. See Fafchamps and El Hamine (2004), and Fafchamps (2004), for an analysis of agglomeration economies in Moroccan manufacturing.

performance that the distinction between quantities and prices is important in general.[3] In our particular case, if for example productivity rises and prices fall in response to increased clustering, the effects of clustering on revenue-based outcome variables (e.g. value-added per worker) would be hard to find. The data set is attractive for other reasons too. The sampling period is 1996-2006, a period during which the number of formal manufacturing firms in Ethiopia grew by 83%, dwarfing net entry rates in most other countries.[4] Since all our regressions are estimated in differences (within firm), such a big change to the enterprise landscape is very useful from an empirical point of view. Further, since we have census data on all firms in the population, we are able to measure local agglomeration more accurately than would be possible with survey data. The panel dimension in the data enables us to control for time invariant unobserved heterogeneity in performance across firms in the regression analysis. This is important, given that firm performance may be correlated with locality-level variables for reasons orthogonal to agglomeration mechanisms. The wide geographical coverage in the data - 82 towns are covered - is another unusual feature compared to other African firm-level data sets, and ensures there is plenty of variation in the data across locations. Finally, the relatively broad coverage of industrial sub-sectors in the data enables us to investigate whether effects of clustering are sector-specific or not.

The rest of the paper is organized as follows. Section 2 discusses the conceptual framework and methodological issues. Section 3 discusses the data and provides summary statistics. The main results are presented in Section 4. Section 5 concludes.

2. Conceptual Framework

Our goal is to establish how firms' output prices and productivity levels are linked to the density of firms clustered in their market. The literature identifies many potential mechansisms. Increased density of firms in the market may result in positive agglomeration effects on productivity. Information spillovers, for example, may be such that proximate firms learn from each other about new technologies, new marketing or management practices, etc. Increased competitive pressure may tend to discipline firms to reduce

[3] See e.g. Melitz (2000), Katayama et al. (2009), and Foster et al. (2008), for a critique of the practice, common in the literature, of inferring productivity effects from analysis of revenue-based outcome variables.

[4] Few African countries undertake industrial censuses. One exception is Ghana. Figures reported by Sandefur (2008) indicate that the number of manufacturing firms with more than 10 workers grew by 23% between 1987 and 2003, thus implying a considerably more modest growth rate than for Ethiopia (see Table 2.1 in Sandefur, 2008).

slack, cut costs and organize production more efficiently (see e.g. Nickell, 1996). Output prices may be affected directly by competition (e.g. equilibrium Cournot prices will fall as more firms compete in the market), or indirectly if productivity effects are passed on to the consumers in the form of price changes. We denote firms, markets and time by i, m and t, respectively, and distinguish three types of own-market firms depending on the distance in product space from a given firm. The most proximate type includes firms that produce the same product r as firm i, in the same market as i, at the same time t; we denote the number of such firms by n_{it}^{r}. The second most proximate firm type constitutes firms present in i's market at time t that do not produce the same product as i but belong to the same industrial sub-sector j (defined at the 2 digit ISIC level) as i; we denote the number of such firms by $n_{it}^{j,-r}$. The least proximate firm type includes firms in the own market at time t belonging to other manufacturing sub-sectors; the number of such firms is denoted by n_{it}^{-j}. Our empirical models are specified as

$$y_{irt} = \alpha_1 n_{it}^{r} + \alpha_2 n_{it}^{j,-r} + \alpha_3 n_{it}^{-j} + \sum_k \beta_k x_{kit} + \eta_i + \rho_r + \sigma_t + \lambda_\tau + \varepsilon_{irt}, \tag{2.1}$$

where y_{irt} denotes the price or productivity associated with product r for firm i time t, and $\alpha_1, \alpha_2, \alpha_3$ and β_k are parameters. Estimating the parameters $\alpha_1, \alpha_2, \alpha_3$ associated with the market density variables is the main goal of the paper. Our baseline specification includes firm fixed-effects (η_i), product fixed-effects (ρ_r), time effects (σ_t), town effects (λ_τ), and a set of observable control variables, denoted x_{kit}. The residual ε_{irt} is assumed orthogonal to the right-hand side variables of (2.1). Some implications of this assumption are discussed below.

Our framework is simple, reasonably flexible, and allows for an extensive set of controls. By distinguishing between different types of own-market firms we can investigate if the composition of firms in the local market matters for firm performance. If the main channel through which changes in the density of firms affect firm-level performance is competitive pressure, one would expect an increase in n_{it}^{r} (same-product firms) to have relatively strong effects on prices and productivity, while changes to $n_{it}^{j,-r}$ and n_{it}^{-j} (same-sector firms, and other firms, respectively) would have weaker effects, since producers of the same product likely pose more of a competitive threat than producers of different products. Along similar lines, if technological spillovers are stronger across firms that produce the same type of product than across firms that produce different products or operate in different industrial sub-sectors, we would have $\alpha_1 > \alpha_2 > \alpha_3$ in the productivity equation.[5]

Several authors, however, have argued that cross-sector externalities may be important. For example, Glaeser et al. (1992) highlight a positive effect of greater density of firms on the "thickness" of the local labor market, reducing the cost of finding the required skills. Jacobs (1969) argued that sectoral diversity raises firm-level productivity via the exchange of information. She also highlighted the possibility that there may be pecuniary externalities across sectors, e.g. due to shared costs for infrastructure.[6] This would be reflected in our model by more important effects of $n_{it}^{j,-r}$ and n_{it}^{-j}.

An important practical issue concerns the definition of a market. Our data set, discussed in Section 3, covers firms in 82 towns across Ethiopia. In our baseline specifications, we define each town as a market, which is reasonable given the high transport costs between towns in the country. In an extension of our analysis, we also allow effects across towns taking distance into acount.

The three enterprise density variables defined above are best interpreted as proxy variables for the underlying driving factors. That is, the number of own-town firms (of a certain type) does not literally measure the extent of externalities or the degree of competition facing firm i, which are unobservable. Alternative proxies for supply density would be possible of course. Our decision to proxy agglomeration and competition by the number of firms is primarily motivated by previous research. Henderson (2003) argues that the number of own-town firms is a good proxy for knowledge spillovers, on the grounds that each firm - rather than each employee - in a locality experiments with the choice of suppliers, inputs etc. Under the hypothesis that knowledge spills over onto other firms in the locality, it is therefore natural to model learning as proportional to the number of firms, not employment. An analogous argument can be constructed regarding competition. We will consider employment based agglomeration variables as part of our robustness checks.

[5] Recent methodolgy used in the literature focuses on productivity distribution to assess the impacts of competition and general agglomeration externalities. Syverson (2007) shows that competition, by driving out inefficient firms, truncates the productivity distribution from below. Whereas Combes et al. (2012) show that the effect of competitive selection can be distinguished from the effect of general agglomeration externailties since the latter right-shift the productivity distribution by improving the productivity of all firms in the cluster.

[6] The empirical evidence on cross-sector externalities (which mainly refers to the USA) is quite mixed. Henderson (1997; 2003) finds that own-sector externalities are stronger than externalities generated by other sectors, while Henderson et al. (1995) and Rosenthal and Strange (2004) report results suggesting that agglomeration effects are not sector-specific. In our framework, this can be tested under the null hypothesis that externalities are not sector-specific, $\alpha_1 = \alpha_2 = \alpha_3$.

There are several studies in the literature investigating the determinants of firm-level performance in Sub-Saharan Africa (e.g. Bigsten et al., 2004; Söderbom and Teal, 2004; Van Biesebroeck, 2005a, 2005b; Frazer, 2005; Bigsten and Gebreeyesus, 2007), but none that makes a clear distinction between prices and quantities at the level of the firm. Recent research has emphasized the importance of this distinction. For example, Katayama et al. (2009) argue that findings that geographically clustered firms are relatively productive may simply reflect high wages and rental costs in urban areas which translates into higher production costs and hence higher output prices, rather than agglomeration economies. More generally, these authors argue that productivity indices based on sales revenues have little to do with technical efficiency, product quality or contributions to social welfare when applied to differentiated product industries, and provide supporting empirical evidence based on Colombian paper producers. A similar argument has been made by Foster et al. (2008; 2012). Thus, inferring productivity effects from revenue-based outcome variables is potentially problematic.

The relationship between competition and firm-level performance in Africa has not been extensively investigated. Harding et al. (2004) find a negative relationship between initial profits and subsequent productivity growth for firms in Ghana, Kenya and Tanzania, which the authors interpret as evidence of a positive effect of competitive pressure on productivity. Aghion et al. (2008) document a negative relationship between lagged sector level price-cost margins and productivity growth among firms in South Africa, which suggests a positive effect of competition on productivity. The outcome variable of interest in these studies, productivity growth, is revenue-based and therefore it is hard to say whether the results reflect effects on prices or physical productivity (or a combination). A richer literature on the relationship between competition and performance exists for other regions, see for example Aghion et al. (2005, 2009), Amiti and Khandelwal (2009), Amiti and Konings (2007), Goldberg et al. (2008, 2010), Khandelwal (2010), and Syverson (2004a, 2004b; 2007).

3. Data and Descriptive Statistics

We use census based panel data on Ethiopian manufacturing firms that employ 10 or more workers and that use power in production. The data set, made available to us by the Central Statistical Agency (CSA) of Ethiopia, contains firm-level data on employment, the book value of the capital stock, raw material expenditures, energy expenditures, as well as detailed information on quantities sold and unit prices for up to 9 products produced by the firm. The fact that this is a census is very useful for our purposes, making it straightforward to compute the agglomeration measures required for the empirical analysis (see Section 2). We have access to data for each year starting in 1996 and ending in 2006, but because the CSA adopted a survey rather than a census design in 2005 we exclude the data for that particular year in our analysis.

As discussed in Section 2, to measure the extent of competition and agglomeration facing firm i in market j at time t, we count the number of firms in the local market (town), distinguishing firms that produce the same product as firm i (n_{it}^{r}), firms that belong to the same sector but do not produce the same product as i ($n_{it}^{j,-r}$), and firms belonging to a different sector than i (n_{it}^{-j}).[7] Figure 1 shows a map of the geographical distribution of firms in the final sample year (2006). Manufacturing firms are present in all the large urban centres of the country, and there is a relatively high concentration of manufacturing production to the capital city (Addis Ababa) and the neighboring areas.

The data set contains detailed information on unit prices, units of measurement, and quantities produced for up to 9 products at the level of the firm. After cleaning the data, we have 14,616 product/firm/year combinations in the period 1996-2006 belonging to 15 two-digit sectors.[8] We can identify a total of 101 different products in the data, these are listed in the Appendix Table 1. One potential concern is that some of these product categories may be too general. For example, the "meat" category likely includes meat of rather varying quality, and it is possible that producers of low quality meat face little direct competition, or learn little, from producers of high quality meat. To test if our results are affected by the inclusion of product categories that may be too broadly defined, we identify a subset of

[7] Firms reporting 'Other product' and firms with missing product information are counted among own-cluster firms producing different products. They will be counted either as part of $n_{it}^{j,-r}$ or n_{it}^{-j}, depending on sector. Multiproduct firms are counted only once, according to the following principle: if a multi-product firm produces the same product as firm i, it will be counted as part of n_{it}^{r}, and not part of $n_{it}^{j,-r}$ or n_{it}^{-j} regardless of what other products it produces; if it does not produce the same product as i but is categorized as belongning to the same sector as i, it will be counted as part of $n_{it}^{j,-r}$ and not n_{it}^{-j}, regardless of whether it also produces products belonging to a different sector.

[8] Several modifications of the raw data were necessary in order to construct the price and output variables. We have standardized the price and unit of measurement for each product, e.g. by expressing all weights in kilograms or tonnes, volumes in liter, area in square meter or square feet depending on the product, etc. We have made corrections in cases where it is obvious that there has been a data entry error, and we have deleted a product category labeled 'other products' from the dataset.

27 products which we believe are less heterogeneous than the other products. We include in this subset products that we consider reasonably homogeneous a priori (e.g. beer, clay bricks, cement, etc.), and products for which the variance of the unit price is reasonably low (on the grounds that low price variance suggests limited quality differences). These products, indicated by * in the Appendix Table 1, account for approximately 7,800 observations in the data, i.e. nearly half of all firm/product/year observations.

Physical productivity is one of our key variables. While we have product-specific data on prices and production volumes, we do not have product-specific data on inputs. That is, if a firm produces two or more products, we would not know how the firm's labor, capital and intermediate inputs have been allocated between these products in the production process. Hence, without further assumptions, we cannot compute a product-specific measure of physical productivity at the level of the firm. Following Foster et al. (2008), who faced the same problem, we therefore impute the input usage for product r using as a weight the share of the sales of product r in the firm's total sales:

$$\theta_{irt} = \frac{P_{irt}Q_{irt}}{\sum_{z} P_{izt}Q_{izt}}. \tag{3.1}$$

Equipped with this weight, we compute product/firm/year-specific measures of physical total factor productivity ($\ln A_{irt}$) assuming a standard Cobb-Douglas production function:

$$\ln A_{irt} = \ln Q_{irt} - \ln F_{irt}, \tag{3.2}$$

where

$$\ln F_{irt} = \alpha_k \ln K_{irt} + \alpha_L \ln L_{irt} + \alpha_M \ln M_{irt} + \alpha_E \ln E_{irt} \tag{3.3}$$

is an aggregate measure of the inputs; and:

$$X_{irt} = \theta_{irt} X_{it},$$

where $X_{it} = \{K_{it}, L_{it}, M_{it}, E_{it}\}$ denotes the respective input observed at the level of the firm in the data, and X_{irt} denotes the level of input assigned to the production of product r. We define α_L, α_M, α_E as the sector averages of the shares of expenditures on these inputs in total sales, and assume constant returns to scale so that $\alpha_K = 1 - \alpha_L - \alpha_M - \alpha_E$. Like many authors of recent productivity papers (e.g. Keller, 2002; Foster et al., 2008; 2012), we prefer this cost-share-based approach for estimating the production function to an econometric approach. One reason is that it has become increasingly clear in recent years that identifying the production function parameters (especially those associated with flexible inputs) by means of an econometric approach requires very strong assumptions (e.g. Bond and Söderbom, 2006; Ackerberg et al., 2007). Another reason is that the debate on what is the best econometric approach for estimating production function parameters appears to be far from settled. Yet another reason is that productivity estimates have turned out to be relatively insensitive to the choice of method (Van Biesebroeck, 2008). Presumably a key reason is that there is considerable variation in the factor inputs across firms, implying that differences in the production function parameters resulting from different estimators will not matter very much for the productivity estimates.

Figure 2 shows how average firm size and the number of firms have evolved over the 1996-2006 period. Average firm size has fallen from 147 employees in 1996 to 104 employees in 2006. The number of firms, however, has grown from 622 to 1,140 over the same period. The fall in average firm size is a result of a small number of large firms having exited and significant entry of small firms. Appendix Table 2 shows how the number of enterprises has developed across sub-sectors. There is considerable variation in the growth rates. The number of firms in the food, leather, metal and furniture sectors has more than doubled over the sampling period, and the number of firms in rubber and plastics has more than tripled. In contrast, there were fewer firms in sectors like footwear, wood and machinery in 2006 than in 1996. There is also notable differences in growth rates across towns (not shown). For example, the growth rate for the number of establishments in Addis Ababa is less than half of that for the rest of the country. Such significant differences in growth rates across sectors and towns is very useful given that all our regressions are estimated with controls for firm fixed effects.

Table 1 shows summary statistics for the variables used in the regression analysis below. Financial variables are expressed in constant 1994/95 Ethiopian Birr. Flow variables are measured on an annual basis while the capital stock variable is defined as the book value of plant and machinery. The average value of log value-added per employee, a crude measure of labor productivity, is equal to 9.35, which corresponds to approximately 11,500 Birr or USD 2,079 per year.[9] The standard deviation log value-added per employee is 1.34 indicating considerable heterogeneity across firms and over time. Average log employment is 3.77 which corresponds to 43 employees. The sample averages for log capital, energy and raw materials per worker correspond to USD 2,700, 134 and 2,800 for per worker capital, energy and raw materials, respectively. One average, a given firm i faces 9.4 firms in the same locality producing the same product as firm i, 25.2 firms in the same locality belonging to the same sector as firm i; and 216.5 firms in the same locality belonging to a different sector. Naturally, these values are heavily influenced by the inclusion of Addis Ababa in the sample. Localities outside the capital city typically host considerably fewer firms. Table 1 also shows summary statistics on the share of new entrants and the (log of) total employment in the own locality and sector. These variables will be used when we consider extensions to the empirical analysis below.

4. Results

4.1. Agglomeration and Output Prices

We start our econometric analysis by investigating the relationship between agglomeration and output prices. Using log price as the dependent variable, we thus re-write (2.1) as:

$$\ln P_{irt} = \alpha_1 n_{it}^{r} + \alpha_2 n_{it}^{j,-r} + \alpha_3 n_{it}^{-j} + \sum_k \beta_k X_{kit} + \eta_i + \rho_r + \sigma_t + \lambda_\tau + \varepsilon_{irt},$$

which is our baseline empirical specification. All results reported below are based on specifications that include controls for firm fixed-effects, product fixed-effects, town fixed-effects and year effects. Standard errors are clustered at the level of the firm throughout, and are thus robust to heteroskedasticity and serial correlation in the error term. We have also estimated the standard errors clustering on firm and town simultaneously (i.e. two-way clustering). The resulting standard errors were usually lower than

[9] The USD exchange rate was 5.53 in January 1995.

those obtained from firm-level clustering. Given that there are only 82 towns in the sample we suspect our two-way clustered errors suffer from small sample bias, and we therefore prefer the (typically higher) firm-level clustered standard errors. For presentational reasons we divide the agglomeration variables by 100 before running the regressions.

Results are shown in Table 2. In column (1) we control for total factor productivity (TFP) defined according to equation (3.2). The results indicate a negative and highly statistically significant effect of agglomeration of same-product firms in the locality on output prices. This is consistent with the hypothesis that an increase in the number of firms in the local area leads to lower prices for the products they produce. The estimated coefficient on n_{it}^{r} implies that the entry of a firm producing the same product as firm i leads to a reduction in the output price charged by firm i by 0.75%.[10] The coefficients on $n_{it}^{j,-r}$ and n_{it}^{-j} are much closer to zero and wholly statistically insignificant. This suggests that the entry of firms that pose a modest or no competitive threat has a small, or no, effect on output prices.

We further obtain a negative and highly significant coefficient on TFP, indicating that productivity improvements result in lower prices. This suggests that firms pass on productivity gains to consumers in the form of lower prices. The point estimate of the coefficient on TFP is equal to -0.22, suggesting that a 10% increase in firm-level productivity is associated with a reduction in the output price of about 2.2%. Clearly, any effect of agglomeration on price operating through productivity - perhaps because of information spillovers - will not be reflected in the coefficients on the agglomeration variables in this regression. We return to the question of whether agglomeration affects productivity directly in the next sub-section.

In columns (2) and (3) we show results for specifications in which we alter the approach for controlling for TFP. In column (2) we remove the TFP variable and add instead physical output ($\ln Q_{irt}$), the aggregate measure of inputs ($\ln F_{irt}$) defined in eq. (3.3), and the intra-firm income share of product j (θ_{irt}; defined in (3.1)), the latter variable included to adjust observed input levels for multiproduct firms. In this specification the coefficient on physical output is interpretable as the effect of physical TFP on

[10] This implies that a one standard deviation increase in n_{it}^{r} (equivalent to 12.6 same-product firms) would lead to a 9.4% drop in output prices. Comparable estimates are hard to find in the literature. Syverson (2007) studies the impact of demand density, measured by the number of construction workers per square mile, on output prices in the US concrete industry. He finds that a one standard deviation increase in density corresponds to a 3.2%, 1.6%, and 6% reduction in the average, median, and maximum market price, respectively.

prices, since a change in output conditional on inputs is driven by a productivity change by definition (eq. (3.2)). The results again indicate a negative and highly statistically significant effect of agglomeration of same-product firms in the locality on output prices. The estimated coefficient on n_{it}^{r} implies that one more competitor leads to a reduction in the price by about 0.65%, i.e. a slightly smaller effect than in column (1). The coefficients on $n_{it}^{j,-r}$ and n_{it}^{-j} are very small and statistically insignificant.

Under the null hypothesis that total factor productivity affects prices and that firm scale has no additional effect on prices, the coefficients on $\ln Q_{irt}$ and $\ln F_{irt}$ in column (2) should be equal in magnitude and opposite in sign. In contrast, if firms with high levels of inputs ("large" firms) charge higher prices than small firms conditional on productivity, this would result in a larger absolute coefficient on $\ln F_{irt}$ than on $\ln Q_{irt}$. The empirical results indicate that the difference between the absolute values of the coefficients on $\ln Q_{irt}$ and $\ln F_{irt}$ is small, suggesting firm size has at most a small effect on prices, conditional on productivity and agglomeration. Column (3) shows regression results based on a specification in which the factor inputs enter separately, which is a more flexible specification than the one shown in column (2). The results are very similar to those in column (2).

In columns (4)-(6) we generalize the specification to investigate if being the only firm of its kind in the locality has a separate effect on prices, over and above the linear agglomeration effects. We thus add to the earlier specifications dummy variables indicating whether firm i is: the only firm producing product j in town; the only firm in the manufacturing sub-sector in town; and the only manufacturing firm in town. For the specification allowing the most general control for TFP (column (6)) we obtain a positive and marginally significant coefficient on the dummy for only firm in this sector. The other coefficients on these dummies are individually insignificant. The total effect of being the only firm in town is obtained by adding up the three coefficients just introduced. This estimated effect is always positive, and significant at the 10% level (but not at the 5% level) throughout. There is thus some evidence that firms without local competitors enjoy a price premium.

4.2. Agglomeration and Physical Productivity

Now consider the effects of enterprise agglomeration on physical productivity. Our baseline specification is as follows:

$$\ln A_{irt} = \alpha_1 n_{it}^{r} + \alpha_2 n_{it}^{j,-r} + \alpha_3 n_{it}^{-j} + \eta_i + \rho_j + \sigma_t + \lambda_\tau + \varepsilon_{irt}, \tag{4.1}$$

where product-level physical productivity ($\ln A_{irt}$) is defined in (3.2). Regression results are shown in Table 3. Similar to the price regressions reported in the previous sub-section, we obtain a significant (at the 5% level) coefficient on the variable measuring the number of same-product firms in the locality (n_{it}^{r}), and insignificant coefficients on the other two agglomeration variables. In our baseline specification, shown in column (1), the coefficient on n_{it}^{r} is estimated at 0.91, indicating that the entry of a firm that produces the same product as firm i increases the product-level physical TFP by 0.91%[11]. The coefficients on $n_{it}^{j,-r}$ and n_{it}^{-j} are smaller in absolute terms and not significantly different from zero. This suggests that firm entry will not result in productivity gains for existing firms unless the entering firm produces the same product as the existing firms. Cross-sectoral agglomeration effects thus seem weak.

In columns (2) and (3) of Table 3 we consider results based on an alternative approach for estimating productivity effects. The specification shown in column (2) is as follows:

$$\begin{aligned} \ln Q_{irt} \quad = \quad & \alpha_1 n_{it}^{r} + \alpha_2 n_{it}^{j,-r} + \alpha_3 n_{it}^{-j} + \beta_1 \ln F_{irt} + \beta_2 \theta_{irt} \\ & + \eta_i + \rho_j + \sigma_t + \lambda_\tau + \varepsilon_{irt}, \end{aligned} \tag{4.2}$$

i.e. rather than subtracting $\ln F_{irt}$ and θ_{irt} from physical output in order to generate total factor productivity ($\ln A_{irt}$), we treat $\ln F_{irt}$ and θ_{irt} as explanatory variables of physical output. This reduces to the baseline specification (4.1) if $\beta_1 = \beta_2 = 1$, which is a testable hypothesis. Results based on (4.2) will thus shed some light on whether the results in column (1) are sensitive to how the productivity measure $\ln A_{irt}$ has been constructed. Reassuringly, even though the coefficients on $\ln F_{irt}$ and θ_{irt} are different from one, the coefficients on the agglomeration variables hardly change at all, compared to column (1). In column (3) we replace $\ln F_{irt}$ by the factor inputs separately. The results are very similar to those in

[11] This implies that a one standard deviation increase in n_{it}^{r} leads to an 11.4% increase in physical TFP.

column (2).

Column (4) shows results for our baseline productivity specification with the dummies for only firm in the three categories added. We obtain negative and statistically significant coefficients on the dummy for only firm in sector. The results imply that if we compare two otherwise identical firms for which there is no other firm in the locality producing the same product, but where there are own-sector firms present in the town for one but not the other, the firm based in the town in which there is neither an own-product nor an own-sector firm has a productivity shortfall of between 14% to 19%. The total effect of being the only firm in town is obtained by adding up the three only-firm coefficients. This estimated effect is always negative, but not significantly different from zero at the 10% level. Thus, overall the evidence for nonlinear productivity effects is weak. The linear effect for own-product firms remains robust.

Comparability of the productivity results with other studies is not straightforward due to differences in the measurment of productivity and agglomeration variables as well as due to differences in level of aggregation.[12] Using a firm-level dataset on Morocan manufacturing firms, Fafchamps and El Hamine (2004) find a negative effect of log of number of firms in own sector and location on revenue-based productivity measure. The authors argue that the negative productivity effect may be due to price effect being compounded with true productivity effect when using value of output as a dependent variable. Syverson (2004) uses a physical productivity measure similar to ours but estimates agglomeration effects on moments of productivity meaures at market level. The main hypothesis advanced by Syverson (2004) is that the entry of producers in a cluster not only drives down mark-ups of incumbent firms, it also forces less efficient firms out of the market. He finds empirical support for the market selection effect using demand (employment) density as exogenous source of variation of producer density for the concrete industry of USA. It is reported that a one standard deviation increase in log demand density implies 2.3% increase in average productivity at market level. Combes et al. (2012), on the other hand, distinguishes between a competition effect and general agglomeration externalities as the latter right-shift the entire productivity distribution whereas competition effect left-truncates the productivity distribution. These authors attribute the main productivity difference between high and low employment density markets in

[12] This is a common problem in agglomeration studies. See for instance, Melo, Graham and Noland (2009) for meta-analysis of estimates of agglomeration economies.

France to agglomeration economies.

4.3. Robustness Checks

In the regressions discussed above we control for firm fixed effects, product fixed effects, town fixed effects, and common time effects. Remaining unobservable determinants of prices and productivity, however, are assumed uncorrelated with the explanatory variables, which may be restrictive. For example, it may be that firms choose to locate in places where productivity or prices have recently started to grow rapidly, since high prices and productivity should result in higher profits. Several aspects of endogenous location are controlled for in the analysis by means of the firm and town fixed effects, however shocks to the incentives to locate in a given town are not controlled for and would go into the time varying residual ε_{irt}. We hypothesize that this would lead to an upward bias in the estimated agglomeration coefficients: high values of ε_{irt}, reflecting positive shocks to prices or productivity, would be associated with a stronger incentive of firms to locate in the area, and hence an increase in the number of firms in that area. This would imply that our estimated effects of agglomeration on output prices are biased towards zero, i.e. the true effect might be a larger negative than what our estimates imply. This would also imply that the estimated effects on physical productivity could be overstated.

Lack of credible instruments implies we cannot allow for endogeneity using an instrumental variables approach. To nevertheless shed some light on how serious the endogeneity problem is likely to be, we re-estimate the baseline price and productivity regressions using lagged instead of contemporaneous values of the agglomeration variables. This should mitigate the endogeneity bias. Because of the gap in the data for 2005 (see Section 3), no lags can be constructed for the 2006 wave which therefore will be dropped altogether.

Results for the price and productivity specifications with the agglomeration variables lagged are shown in Table 4, columns (1) and (2), respectively. As a result of lagging the explanatory variables we lose 5,916 observations or about 40% of our sample. The estimated price and productivity effects of agglomeration are nevertheless similar to those shown above. The coefficient on n_{it}^{r} is estimated at -0.65 in the price regression and 0.93 in the productivity regression. The standard errors are somewhat higher than in previous specifications, but the coefficients are still significant at the 10% level or better. The estimated coefficients on the other two agglomeration variables are close to zero and statistically insignificant. The coefficient on TFP in the price equation remains negative and highly statistically significant.

Next we investigate if the agglomeration effects depend on the heterogeneity of products within the different product categories. As discussed in Section 3, we have identified 27 products in the data that, in our view, are more homogeneous than the other product categories (see starred products in Appendix Table 1). We define a dummy variable, $HOMPROD_{irt}$, interact this with the agglomeration variables, and add the interaction terms to the baseline price and productivity specifications. Under the null hypothesis that the agglomeration effects do not vary across products of differing heterogeneity, the coefficients on these interaction terms are equal to zero. Results, shown in columns (3) and (4) in Table 4, are consistent with this null hypothesis: in no case do we obtain a significant coefficient on the interaction terms. We infer from these results that product heterogeneity within categories, if present, does not pose a serious problem given our purposes.

Thus far our agglomeration and competition variables have been based on the number of firms in the relevant locality. This measure does not take into account differences in firm size within towns, which may be an important omission. For example, it could be that the extent of information spillovers depends on the number of individuals associated with manufacturing production, rather than (or in addition to) the number of firms. Moreover, our procedure implies that towns with a few large firms would be coded as smaller agglomerations than towns with many small firms, even though the scale of production may be much larger in towns with large firms. In his analysis of agglomeration effects in US manufacturing, Henderson (2003) alternates between using the number of firms and total employment in the towns as measures of agglomeration. We now investigate if using the number of employees in the locality changes any of our main findings. We cannot assign employees to specific products within firms, so we cannot construct the analogue of n_{it}^{r} based on employment. We therefore distinguish only between own-sector employees and all employees in the own town.

Columns (5) and (6) in Table 4 shows regression results for the baseline models with total own-sector employment, and total employment in different sectors, in the own town entered as additional regressors (both in logs).[13] In both models the employment based agglomeration variables are insignificant. Adding them to the specification has very small effects on the point estimates of the coefficients on n_{it}^r. Adding controls for sector specific time trends has similarly small effects on the coefficients of interest (results not shown).

4.4. Effects across Localities

As is clear from the geographical distribution of firms shown in Figure 1, some towns in our data set are located close to each other. It is conceivable that productivity and price effects diffuse across, as well as within, towns. The large size of Ethiopia combined with the poor infrastructure imply high transport costs. We therefore suspect that, if there are spillover effects across towns, these are probably limited in scope and dependent on the physical distance between towns. We now use data on the geographic coordinates of each town and test for agglomeration effects, as defined previously, across towns.

We have considered three ways of measuring effects across towns: by counting the number of firms in the nearest town at a particular point in time; by counting the number of firms within a 100 kilometer radius at a particular point in time; and by computing a weighted sum of all firms in the country at a particular point in time using the inverse of the distance between towns as the weight. We add these variables to the baseline specifications analyzed above. Overall we have found only weak evidence for agglomeration effects across towns. Table 5 shows results for specifications in which cross-town effects are tested for using the number of firms in the nearest town.[14] The dependent variable is log price in columns (1)-(2) and TFP in columns (3)-(4). The results in columns (1) and (3) suggest the effect of the number of firms in the nearest town is negative on prices and positive on productivity. Compared to the estimated effects of the number of own-product firms in the same locality, the cross-town effects are much smaller. In columns (2) and (4) we distinguish between firms in the nearest town producing the same product as firm i and other firms in the nearest town. The coefficients on own-product firms in the nearest town is very imprecisely estimated and not significantly different from zero, while the coefficients on other firms in the nearest town are negative and positive, respectively, in the price and productivity

[13] The firm's own employment is excluded when computing these employment-based agglomeration variables.
[14] Results for the other specifications are available on request. These provide no strong support for cross-town effects.

regressions.

We tentatively conclude from this analysis that the number of firms in the nearest town may impact prices and productivity in the own town, but these effects appear to be small and become weaker still once we go beyond the nearest neighbor. Moreover, we note that our main results are robust: the number of firms producing the same product as the own firm has a negative and highly significant effect on own output prices, and a positive and highly significant effect on own productivity in these extended specifications.

4.5. Tests for Heterogenous Effects

Finally we report results from two tests for heterogenous agglomeration effects. Columns (1)-(2) of Table 6 shows results for specifications in which we have added interaction terms between the agglomeration variables and the ratio of own-sector entrants in the town to incumbents in the relevant category of firms.[15] If our cluster size variables are indeed picking up agglomeration and competition effects, we would expect these effects to be driven by increased entry rather than by exit. Hence, we expect the agglomeration effects to be particularly strong in environments in which the share of new entrants is large. The results are broadly consistent with this expectation. In the price regression, shown in column (1), we obtain a negative and significant coefficient on the interaction term between the number of firms producing the same product as firm i and the share of new entrants. This implies that the higher the share of new entrants, the larger is the agglomeration effect in absolute terms. In this specification, the coefficients on the non-interacted agglomeration variables are interpretable as the effects in environments in which there is no new entry. The results suggest that if the share of new entrants is zero, there are no agglomeration effects. In other words, if the only reason the number of firms changes is that some firms have exited, there will no effect on output prices. A counter-intuitive result in column (1) is the positive and significant coefficient on the interaction between the share of new entrants and the number of firms in a different sector in the town. Quantitatively, however, this effect is never particularly important.

Similar results are obtained for the productivity equation (column (2)). The coefficient on the inter-

[15] Since the number of own-product firms is typically small, entry rates for own-product firms become rather noisy. For our present purposes, we prefer entry rates defined at the sector-town level.

action term between the number of firms producing the same product as firm i and the share of new entrants is positive and statistically significant at the 5% level. This implies that the higher the share of new entrants, the larger is the agglomeration effect on productivity. If the share of new entrants is zero, a change in the number of own-product firms in the town does not have a significant effect on productivity.

Finally, we focus on our underlying assumption that firms operate in localized markets. If agglomeration effects arise because new entry into a cluster creates competitive pressure and reduces market share of existing firms, these effects should be less important for firms that are not restricted to selling their products in local markets. We test this hypothesis by interacting the agglomeration variables with the share of exporters in the relevant group of firms. If the share of exporters is high we expect the agglomeration effects to be less strong, since markets populated by exporters will tend to be less localized. The coefficients on the non-interacted agglomeration variables are now interpretable as the effects in environments in which there are no exporters; we expect these to be stronger than previously since such markets are more localized. The results, shown in columns (3) and (4) of Table 6, suggest that the coefficients on n_{it}^{r} are larger than in the baseline specifications. This is consistent with the hypothesis that agglomeration effects on prices and productivity are strongest in environments in which no firms export. However the difference compared to the baseline specification is small, and the coefficients on the export interaction terms are mostly insignificant.

5. Conclusions

In this paper we have used census panel data on Ethiopian manufacturing firms to empirically analyze the effects of enterprise clustering on two key determinants of firm performance: output prices and physical productivity. We show that distinguishing between productivity and prices is crucial for understanding the effects of agglomeration.

We find a negative and statistically significant effect of agglomeration of own-product firms on prices, suggesting that new entry leads to higher competitive pressure in the local economy. All else equal, this is positive for consumer welfare but negative for enterprise revenues. In addition, we find a positive and statistically significant effect of the agglomeration of own-product firms on physical productivity, consistent with the notion that clustering leads to positive externalities. All else equal, this is positive both for consumer welfare and for enterprise revenues.

Our findings thus suggest there is a lot to be said for encouraging local competition and agglomeration of firms: individual firms will see their productivity rise and their profit margins reduced, and both effects benefit Ethiopian consumers. However, these effects arise only if the agglomerating firms overlap in terms of the product they produce. Across firms that produce different products, we find no statistically significant relationship between agglomeration and firm-level output prices and productivity. Moreover, the negative price effects suggest that firms may not have strong incentives to agglomerate endogenously. This relates to a broader question as to why, if agglomeration externalities are so important, do we not see more agglomeration of firms in Sub-Saharan Africa. A popular response is that there are coordination problems and policy can help overcome these (e.g. Page, 2012). Our findings suggest we should look more closely at the incentives of firms to form clusters endogenously – taking into account that firms may weigh externality gains against the adverse effects of stronger competition on prices and revenues. Market structure and integration may play an important role in this context: if markets are localized, local rents may be available and therefore solving the coordination problem may not be enough; but if markets are competitive and integrated, firms cannot avoid competition by strategic location, which may strengthen their incentives to agglomerate. These appear to be interesting questions for future research.

A common argument in the discussion of industrial development in poor countries is that agglomeration can be a source of improved economic performance (e.g. Collier, 2007; Page, 2012). For Africa, little empirical research exists on the links between agglomeration, productivity and prices. Our empirical approach is neither experimental nor structural, and it will remain an open question as to whether our results can be given a causal interpretation. The premise of our analysis is that, given how little is known about agglomeration mechanisms in Africa, a transparent reduced form approach is a natural way of starting to put together quantitative evidence. Moreover, as shall become clear below, the subject does not lend itself easily to an experimental analysis, except perhaps in a lab setting.

References

[1] Ackerberg, D., C. Lanier Benkard, S. Berry, and A. Pakes. "Econometric Tools for Analyzing Market Outcomes." *Handbook of Econometrics* 6 (2007): 4171–4276.

[2] Aghion, P., R. Blundell, R. Griffith, P. Howitt, and S. Prantl. "The Effects of Entry on Incumbent Innovation and Productivity." *Review of Economics and Statistics* 91, 1 (2009): 20–32.

[3] Aghion, P., M. Braun, and J. Fedderke. "Competition and Productivity Growth in South Africa." *Economics of Transition* 16, 4 (2008): 741–768.

[4] Aghion, P., N. Bloom, R. Blundell, R. Griffith, and P. Howitt. "Competition and Innovation: An Inverted-U Relationship." *Quarterly Journal of Economics* 120, 2 (2005): 701 –728.

[5] Amiti, M., and A. K. Khandelwal. "Import Competition and Quality Upgrading." Working Paper 15503. Cambridge, MA: NBER (2009).

[6] Amiti, M., and J. Konings. "Trade Liberalization, Intermediate Inputs, and Productivity: Evidence from Indonesia." *American Economic Review* 97, 5 (2007): 1611-1638.

[7] Bigsten, A., P. Collier, S. Dercon, M. Fafchamps, B. Gauthier, J. W. Gunning, A. Oduro, R. Oostendorp, C. Pattillo, M. Söderbom, F. Teal, and A. Zeufack. "Do African Manufacturing Firms Learn from Exporting?" *Journal of Development Studies* 40, 3 (2004): 115–141.

[8] Bigsten, A., and M. Gebreeyesus. "The Small, the Young, and the Productive: Determinants of Manufacturing Firm Growth in Ethiopia." *Economic Development and Cultural Change* 55, 4 (2007): 813–840.

[9] Blien, U., J. Suedekum, and K. Wolf. "Local Employment Growth in West Germany: A Dynamic Panel Approach." *Labour Economics* 13, 4 (2006): 445–458.

[10] Bond, S., and M. Söderbom. "Identification and Estimation of Cobb-Douglas Production Parameters from Micro Data," Mimeo. University of Oxford, The Department of Economics (2006).

[11] Collier, P. *The Bottom Billion. Why the Poorest Countries are Failing and What Can Be Done About It.* New York: Oxford University Press (2007).

[12] Combes, P. P. "Economic Structure and Local Growth: France, 1984-1993." *Journal of Urban Economics* 47, 3 (2000): 329–355.

[13] Combes, P. P., G. Duranton, L. Gobillon, D. Puga, and S. Roux. "The Productivity Advantages of Large Cities: Distinguishing Agglomeration from Firm Selection." IZA Discussion Paper No. 6502 (2012).

[14] De Lucio, J. J., J. A. Herce, and A. Goicolea. "The Effects of Externalities on Productivity Growth in Spanish Industry." *Regional Science and Urban Economics* 32, 2 (2002): 241–258.

[15] Fafchamps, M. "Manufacturing Growth and Agglomeration Effects." Working Paper 204-33. Oxford: University of Oxford, CSAE (2004).

[16] Fafchamps, M., and S. El Hamine. "Firm Productivity, Wages, and Agglomeration Externalities." Working Paper 204-32. Oxford: University of Oxford, CSAE (2004).

[17] Fafchamps, M., and M. Söderbom. "Network Proximity and Business Practices in African Manufacturing." *World Bank Economic Review*, forthcoming (2013).

[18] Foster, L., J. C. Haltiwanger, and C. Syverson. "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?" *American Economic Review*, 98, 1 (2008): 394-425.

[19] Foster, L., J. C. Haltiwanger, and C. Syverson. "The Slow Growth of New Plants: Learning About Demand?" Working Paper 17853, Cambridge, MA: NBER (2012).

[20] Frazer, G. "Which Firms Die? A Look at Manufacturing Firm Exit in Ghana." *Economic Development and Cultural Change* 53, 3 (2005): 585–617.

[21] Glaeser, E. L., H. D. Kallal, J. A. Scheinkman, and A. Shleifer. "Growth in Cities." *Journal of Political Economy* 100, 6 (1992): 1126–1152.

[22] Goldberg, P. K., A. Khandelwal, N. Pavcnik, and P. Topalova. "Imported Intermediate Inputs and Domestic Product Growth: Evidence from India." Working Paper 14416. Cambridge, MA: NBER (2008).

[23] Goldberg, P. K., A. K. Khandelwal, N. Pavcnik, and P. Topalova. "Multiproduct Firms and Product Turnover in the Developing World: Evidence from India." *Review of Economics and Statistics* 92, 4 (2010): 1042–1049.

[24] Harding, A., M. Söderbom, and F. Teal. "Survival and Success Among African Manufacturing Firms." Working Paper WPS/2004-05. Oxford: University of Oxford, CSAE (2004).

[25] Henderson, V. "Externalities and Industrial Development." *Journal of Urban Economics* 42, 3 (1997): 449–470.

[26] Henderson, J. V. "Marshall's Scale Economies." *Journal of Urban Economics* 53, 1 (2003): 1–28.

[27] Henderson, J. V., A. Kuncoro, and M. Turner. "Industrial Development in Cities." *Journal of Political Economy* 103, 5 (1995): 1067–1090.

[28] Jacobs, J. *The Economy of Cities*. New York: Random House (1969).

[29] Katayama, H., S. Lu, and J. R. Tybout. "Firm-level Productivity Studies: Illusions and a Solution." *International Journal of Industrial Organization* 27, 3 (2009): 403–413.

[30] Keller, W. "Geographic Localization of International Technology Diffusion." *American Economic Review* 92, 1 (2002): 120-142.

[31] Khandelwal, A. "The Long and Short (of) Quality Ladders." *Review of Economic Studies* 77, 4 (2010): 1450–1476.

[32] Marshall, A. *Principles of Economics*. Guillebaud Edition. London: Macmillan (1920).

[33] Melitz, M. J. "Estimating Firm-level Productivity in Differentiated Product Industries." Mimeo. Department of Economics, Harvard University (2000).

[34] Melo, P. C., D. J. Graham, and R. B. Noland. "A Meta-analysis of Estimates of Urban Agglomeration Economies." *Regional Science and Urban Economics* 39, 3 (2009): 332–342.

[35] Nickell, S. J. "Competition and Corporate Performance." *Journal of Political Economy* 104, 4 (1996): 724–746.

[36] Page, J. "Can Africa Industrialise?" *Journal of African Economies* 21, suppl 2 (2012): ii86–ii124.

[37] Porter, M. E. *The Competitive Advantage of Nations:With a New Introduction*. New York: The Free Press (1990).

[38] Rosenthal, S. S., and W. C. Strange. "Evidence on the Nature and Sources of Agglomeration Economies." *Handbook of Regional and Urban Economics* 4 (2004): 2119–2171.

[39] Sandefur, J. "On the Evolution of the Firm Size Distribution in an African Economy." in J. Sandefur. *Essays on Labour and Credit Markets in Africa.* Oxford: University of Oxford (2008).

[40] Schmitz, H., and K. Nadvi. "Clustering and Industrialization: Introduction." *World Development* 27, 9 (1999): 1503–1514.

[41] Sonobe, T., and K. Otsuka. *Cluster-Based Industrial Development. An East Asian Model*. Basingstoke: Palgrave McMillan (2006).

[42] Sonobe, T., and K. Otsuka. *Cluster-Based Industrial Development: A Comparative Study of Asia and Africa.* London: Palgrave McMillan (2011).

[43] Swann, G.M.P., M. Prevezer, and D. Stout (eds). *The Dynamics of Industrial Clustering: International Comparisons in Computing and Biotechnology*. Oxford: Oxford University Press (1998).

[44] Söderbom, M., and F. Teal. "Size and Efficiency in African Manufacturing Firms: Evidence from Firm-level Panel Data." *Journal of Development Economics* 73, 1 (2004): 369–394.

[45] Syverson, C. "Product Substitutability and Productivity Dispersion." *Review of Economics and Statistics* 86, 2 (2004a): 534–550.

[46] Syverson, C. “Market Structure and Productivity: A Concrete Example.” *Journal of Political Economy* 112, 6 (2004b): 1181–1222.

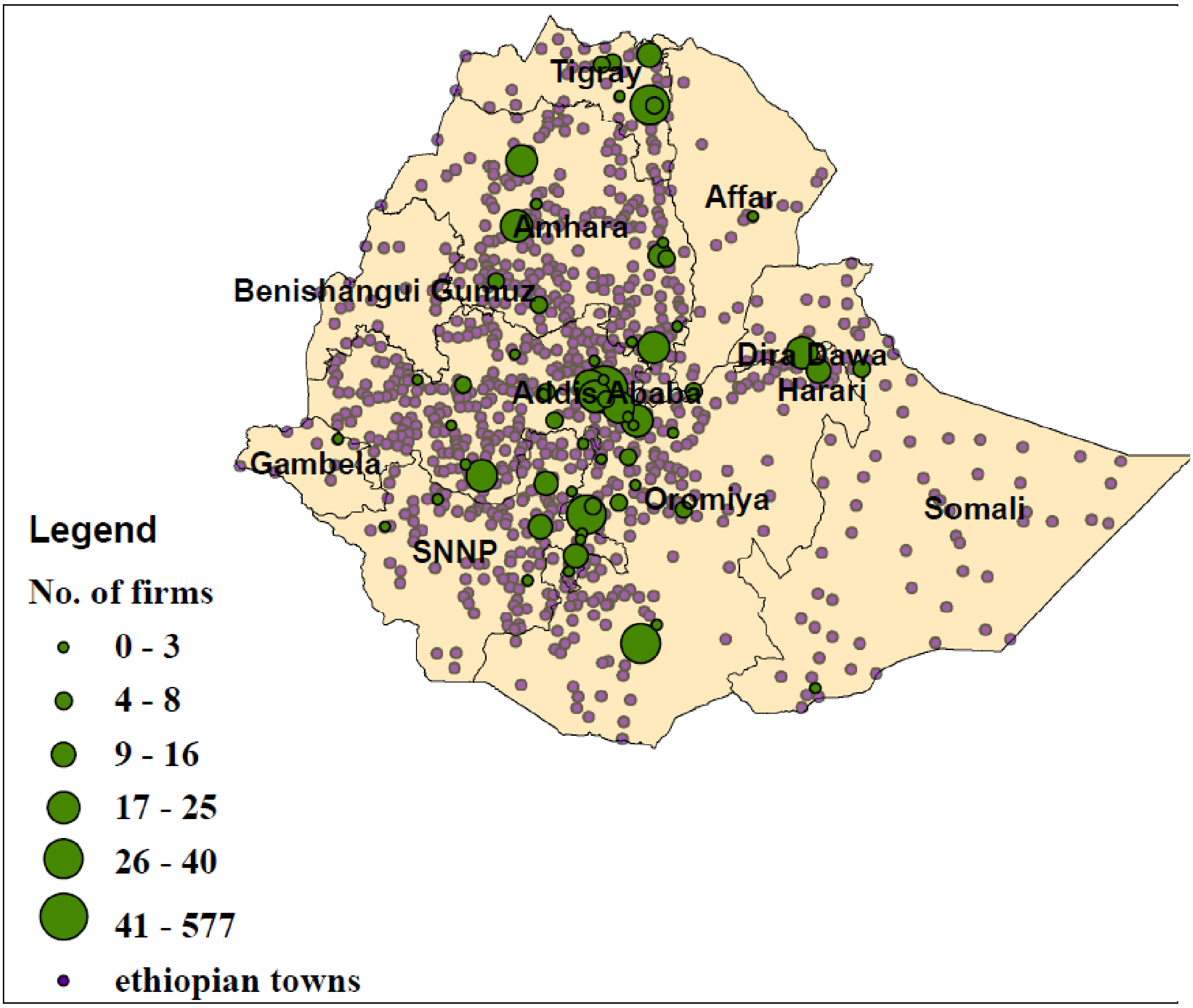
[47] Syverson, C. “Prices, Spatial Competition and Heterogenous Producers: An Empirical Test.” *Journal of Industrial Economics* 55, 2 (2007): 197–222.

[48] Van Biesebroeck, J. “Exporting Raises Productivity in sub-Saharan African Manufacturing Firms.” *Journal of International Economics* 67, 2 (2005a): 373–391.

[49] Van Biesebroeck, J. “Firm Size Matters: Growth and Productivity Growth in African Manufacturing.” *Economic Development and Cultural Change* 53, 3 (2005b): 545–583.

[50] Van Biesebroeck, J. “The Sensitivity of Productivity Estimates.” *Journal of Business and Economic Statistics* 26, 3 (2008): 311–328.

Figure 1: The Geographical Distribution of Firms 2005/6



Note: The green filled circles indicate locations in which at least one firm was located in 2005/6. The size of the circles indicates the number of firms located in that town (see graph legend).

Figure 2: Firm and Employment Trends, 1996-2006

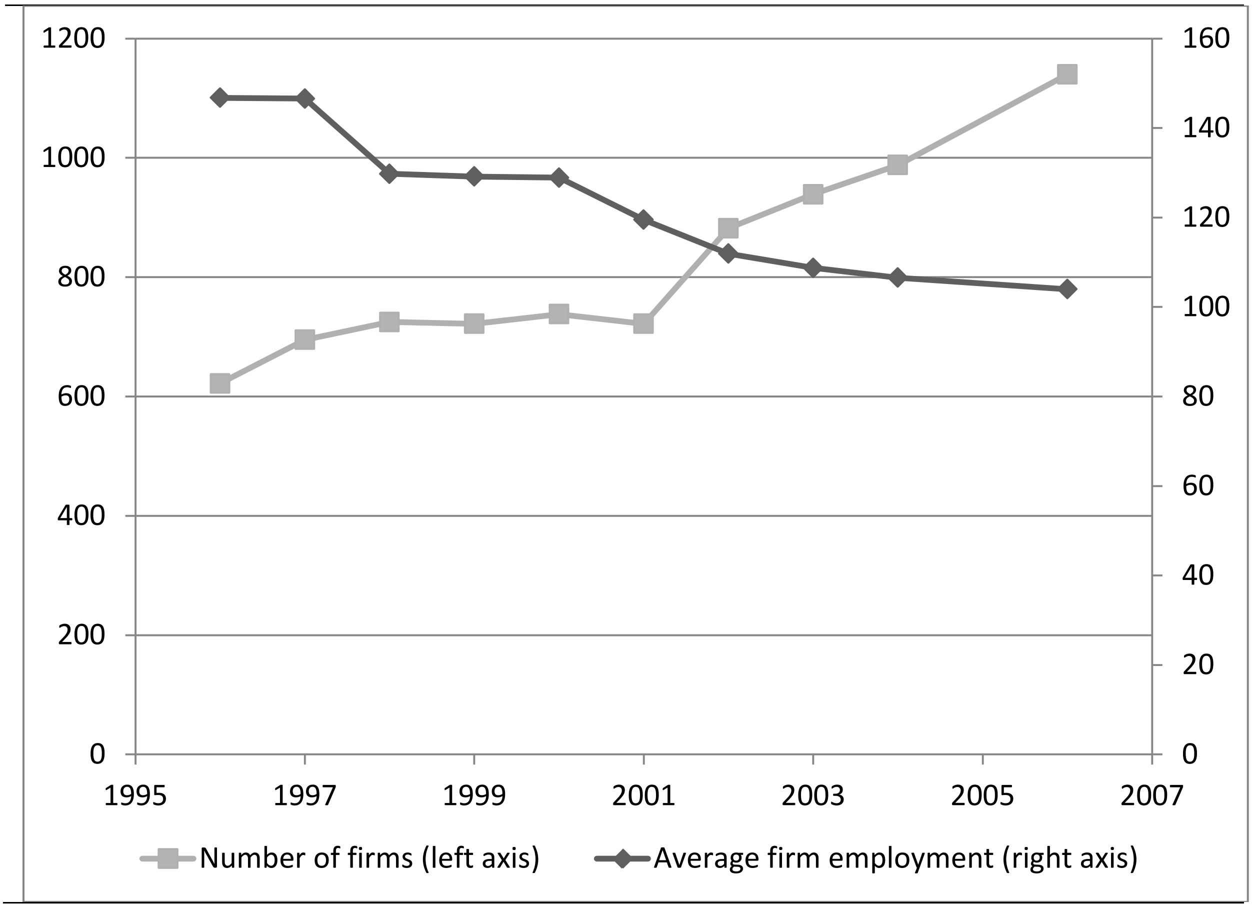


Table 1: Summary Statistics

	Mean	Std. Dev.
log Value-added per employee	9.35	1.34
log Employment	3.77	1.39
log Capital stock per employee	9.65	1.94
log Energy expenditures per employee	6.61	1.55
log Raw material expenditures per employee	9.61	1.47
Number of firms in the same town and sector producing the same product as firm i, at time t (n_{it}^{r})	9.41	12.57
Number of firms in the same town and sector as firm i not producing the same product as firm i, at time t ($n_{it}^{j,-r}$)	25.24	35.64
Number of firms in the same town as firm i that belong to a different sector than i, at time t (n_{it}^{-j})	216.46	201.16
Share of new entrants in own town, at time t	0.18	0.15
Share of new entrants in own town and sector, at time t	0.17	0.21
Share of exporters in own town, at time t	0.05	0.12
Share of exporters in own town and sector, at time t	0.06	0.20
log Total employment in own town and same sector, at time t	8.49	2.80
log Total employment in own town and different sector, at time t	8.20	3.06
Observations	4858	
Firms	1341	

Note: Financial variables are expressed in constant 1994/95 Ethiopian Birr. The exchange rate to the USD in January 1995 was 5.53.

Table 2: Agglomeration and Output Prices

	(1)	(2)	(3)	(4)	(5)	(6)
n_{it}^{r} / 100	-0.754***	-0.648***	-0.685***	-0.756***	-0.649***	-0.688***
	(0.236)	(0.220)	(0.219)	(0.236)	(0.220)	(0.219)
$n_{it}^{j,-r}$ / 100	-0.0431	0.00996	-0.0483	-0.0524	0.00313	-0.0574
	(0.131)	(0.130)	(0.130)	(0.133)	(0.132)	(0.132)
n_{it}^{-j}/ 100	0.0149	0.00489	0.0171	0.00924	-0.00100	0.0108
	(0.0353)	(0.0335)	(0.0324)	(0.0353)	(0.0334)	(0.0323)
TFP	-0.220***			-0.220***		
	(0.0203)			(0.0203)		
Quantity of output		-0.277***	-0.285***		-0.277***	-0.285***
		(0.0197)	(0.0200)		(0.0197)	(0.0200)
All inputs		0.334***			0.334***	
		(0.0275)			(0.0275)	
log Capital stock			0.00506			0.00392
			(0.0126)			(0.0125)
log Employment			0.0543**			0.0566**
			(0.0249)			(0.0245)
log Energy			0.0115			0.0113
			(0.00939)			(0.00938)
log Raw materials			0.221***			0.221***
			(0.0176)			(0.0175)
Only firm with this product in town				-0.0302	-0.0177	-0.0276
				(0.0422)	(0.0418)	(0.0417)
Only firm in this sector in town				0.0777	0.0828	0.0934*
				(0.0520)	(0.0516)	(0.0512)
Only firm in town				0.187	0.162	0.168
				(0.126)	(0.114)	(0.119)
Revenue share		0.144***	0.152***		0.144***	0.152***
		(0.0177)	(0.0180)		(0.0177)	(0.0180)
Observations	14161	14161	14161	14161	14161	14161
Firms	1176	1176	1176	1176	1176	1176

Note: The dependent variable is log price. All regressions include year dummies, town dummies, product dummies, and controls for firm fixed effects. Standard errors are clustered at the level of the firm. Significance at the 10, 5 and 1 percent level is indicated by *, ** and ***, respectively.

Table 3: Agglomeration and Productivity

	(1) TFP	(2) log Output	(3) log Output	(4) TFP	(5) log Output	(6) log Output
n_{it}^{r} / 100	0.911**	0.922**	0.837**	0.915**	0.926**	0.839**
	(0.394)	(0.394)	(0.388)	(0.394)	(0.395)	(0.388)
$n_{it}^{j,-r}$ / 100	0.0346	0.0968	-0.0548	0.0455	0.107	-0.0498
	(0.222)	(0.224)	(0.222)	(0.225)	(0.227)	(0.223)
n_{it}^{-j} / 100	-0.0383	-0.0472	-0.0124	-0.0274	-0.0369	-0.00382
	(0.0535)	(0.0532)	(0.0515)	(0.0538)	(0.0536)	(0.0518)
All inputs		1.084***			1.082***	
		(0.0389)			(0.0388)	
log Capital stock			0.0170			0.0177
			(0.0224)			(0.0224)
log Employment			0.151***			0.149***
			(0.0477)			(0.0474)
log Energy			0.0819***			0.0822***
			(0.0172)			(0.0172)
log Raw materials			0.676***			0.675***
			(0.0275)			(0.0274)
Only firm with this product in town				0.0451	0.0433	0.0113
				(0.0538)	(0.0533)	(0.0525)
Only firm in this sector in town				-0.188**	-0.173*	-0.140*
				(0.0874)	(0.0890)	(0.0824)
Only firm in town				-0.0666	-0.0819	-0.0707
				(0.156)	(0.157)	(0.145)
Revenue share		0.965***	0.966***		0.965***	0.966***
		(0.00841)	(0.00841)		(0.00841)	(0.00841)
Observations	14161	14161	14161	14161	14161	14161
Firms	1176	1176	1176	1176	1176	1176

Note: All regressions include year dummies, town dummies, product dummies, and controls for firm fixed effects. Standard errors are clustered at the level of the firm. Significance at the 10, 5 and 1 percent level is indicated by *, ** and ***, respectively.

Table 4: Robustness to Lag Effects, Product Type Interaction Effects, and Employment Based Measures of Agglomeration

	Lagged agglomeration variables		Interaction terms: Product type x agglomeration variables		Cluster employment	
	(1)	(2)	(3)	(4)	(5)	(6)
	log price	TFP	log price	TFP	log price	TFP
n_{it}^{r} / 100	-0.653*	0.930**	-0.826***	0.838*	-0.748***	0.909**
	(0.368)	(0.425)	(0.279)	(0.461)	(0.236)	(0.394)
$n_{it}^{j,-r}$ / 100	-0.0331	0.0696	-0.0901	0.138	-0.0414	0.0434
	(0.263)	(0.297)	(0.153)	(0.248)	(0.131)	(0.222)
n_{it}^{-j} / 100	-0.0546	0.0103	0.0292	-0.0416	0.0141	-0.0395
	(0.0567)	(0.0633)	(0.0395)	(0.0569)	(0.0356)	(0.0538)
TFP	-0.181***		-0.221***		-0.220***	
	(0.0259)		(0.0203)		(0.0204)	
Homogeneous product x			0.283	0.272		
n_{it}^{r} / 100			(0.328)	(0.481)		
Homogeneous product x			0.111	-0.205		
$n_{it}^{j,-r}$ / 100			(0.135)	(0.168)		
Homogeneous product x			-0.0241	0.00780		
n_{it}^{-j} / 100			(0.0270)	(0.0317)		
Number of employees in own					-0.0485	0.0605
town & own sector					(0.0313)	(0.0391)
Number of employees in own					0.00905	0.00998
town & different sector					(0.0223)	(0.0222)
Observations	8245	8245	14161	14161	14161	14161
Firms	699	699	1177	1177	1176	1176

Note: All regressions include year dummies, town dummies, product dummies, and controls for firm fixed effects. Standard errors are clustered at the level of the firm. Significance at the 10, 5 and 1 percent level is indicated by *, ** and ***, respectively. The variables n_{it}^{r}, $n_{it}^{j,-r}$ and n_{it}^{-j} are lagged in columns (1)-(2).

Table 5: Tests for Agglomeration Effects across Towns

	Dependent variable: log price		Dependent variable: TFP	
	(1)	(2)	(3)	(4)
n_{it}^{r} / 100	-0.768***	-0.761***	0.935**	0.946**
	(0.236)	(0.236)	(0.394)	(0.395)
$n_{it}^{j,-r}$ / 100	-0.0466	-0.0496	0.0464	0.0420
	(0.132)	(0.132)	(0.223)	(0.222)
n_{it}^{-j} / 100	0.0111	0.0126	-0.0241	-0.0219
	(0.0363)	(0.0363)	(0.0542)	(0.0544)
TFP	-0.220***	-0.221***		
	(0.0204)	(0.0204)		
Number of firms in nearest neighboring	-0.0991*		0.186*	
town /100	(0.0513)		(0.100)	
Number of firms in nearest neighboring		1.239		2.158
town producing the same product /100		(1.811)		(2.027)
Number of firms in nearest neighboring		-0.113**		0.166*
town producing different product /100		(0.0539)		(0.100)
Observations	14161	14161	14161	14161
Firms	1176	1176	1176	1176

Note: All regressions include year dummies, town dummies, product dummies, and controls for firm fixed effects. Standard errors are clustered at the level of the firm. Significance at the 10, 5 and 1 percent level is indicated by *, ** and ***, respectively.

Table 6: Tests for Heterogeneous Agglomeration Effects Depending on the Share of New Entrants and Exporters in Own Town and Sector

	(1) log price	(2) TFP	(3) log price	(4) TFP
n_{it}^{r} / 100	-0.386	0.327	-0.810***	1.066***
	(0.309)	(0.547)	(0.253)	(0.391)
$n_{it}^{j,-r}$ / 100	0.0821	0.0234	0.0894	0.109
	(0.152)	(0.224)	(0.161)	(0.236)
n_{it}^{-j} / 100	-0.0383	0.000831	0.0145	-0.0141
	(0.0357)	(0.0555)	(0.0366)	(0.0555)
TFP	-0.220***		-0.221***	
	(0.0202)		(0.0205)	
Share of new entrants in own sector and town	-0.00283	-0.142		
	(0.0454)	(0.0885)		
Share of new entrants in own sector and town x n_{it}^{r}/ 100	-2.094*	2.968**		
	(1.084)	(1.485)		
Share of new entrants in own sector and town x $n_{it}^{j,-r}$ / 100	-0.891**	0.313		
	(0.371)	(0.505)		
Share of new entrants in own sector and town x n_{it}^{-j} / 100	0.253***	-0.113		
	(0.0543)	(0.0863)		
Share of exporters in own sector and town			-0.0467	0.254**
			(0.110)	(0.122)
Share of exporters in own sector and town x n_{it}^{r}/ 100			3.384	-7.256
			(4.705)	(6.189)
Share of exporters in own sector and town x $n_{it}^{j,-r}$ / 100			-4.162	-2.396
			(2.773)	(2.849)
Share of exporters in own sector and town x n_{it}^{-j} / 100			0.0327	-0.0977
			(0.0878)	(0.0913)
Observations	14161	14161	14161	14161
Firms	1176	1176	1176	1176

APPENDIX

Table A1: Products in the dataset

Acrylic (yarn)	Cotton yarn*	Liquor*	Raw cotton
Alchohol (non-drink)	Crown cork	Macaroni and pasta	Semiprocessed skins
Animal feed	Crust hides and wetblue hides*	Malt	Sewing thread
Antibiotics	Edible oil*	Marble	Shirts
Ballpen	Electric wires	Marmalade	Soap
Bed sheet	Embroidery thread	Meat	Sugar*
Beer*	Fafa, dube, edget,metin	Metalic door	Sweater
Biscuts and cakes (excl. galleta)	Flour (other)	Metalic window	Sweets (candy)
Blankets	Flour (wheat)*	Milk pasterurized	Syrup
Boxing paper	Foam	Mineral water	Tablets
Bread*	Galleta	Molasses	Tea*
Bricks of clay*	Glass bottles	Motor vehicles spring	Timber*
Butter and ghee	Glasses	Nails*	Tomato paste
Candles	Gravel*	Nylon fabrics*	Tubes
Canvas and rubber shoes	Gunny bags	Oil cakes*	Tyres
Capsules	Hosieries	Ointment	Varnishes and lacquers
Carbon dioxide	Injection of 100A	Orange juce	Vasilin*
Carpets	Iron bars	Oxygen	Vegetable soup
Cement*	Iron sheets	Paints	Wearing apparel (excl. leather)
Cement blocks*	Jano thread	Palstic crate	Wearing apparel (leather)
Cement floor tiles*	Leather garment*	Paper	Wine
Cement tubes	Leather shoes and boots*	Paraffine*	Wires*
Cheese	Leather sole	Particle board	Zign and shiro wet (minchet abish)
Cigarettes	Leather upper and lining	Plastic footwear*	
Coffe (Milled)*	Lemonade (soft drinks)*	Plastic sole	
Cotton fabrics*	Lime	Polyethylene products	

*Note:** indicates products used to check robustness of the main results to product homogeneity

Table A2 Growth in the Number of Establishments by Sector

Year	Food	Beverage	Textiles	Apparel	Leather	Footwear	Wood	Paper&printing
1996	139	21	32	23	8	55	26	43
1997	154	22	33	26	11	49	23	45
1998	185	21	32	26	11	46	16	53
1999	186	23	33	27	14	35	16	55
2000	189	25	33	25	15	37	16	63
2001	194	26	31	24	15	37	14	51
2002	239	28	34	29	14	38	21	73
2003	245	28	36	32	15	41	17	73
2004	265	29	38	35	17	45	20	72
2005	183	31	38	27	17	43	18	77
2006	299	35	42	30	17	43	21	85
Growth 96-06	115%	67%	31%	30%	113%	-22%	-19%	98%

Year	Chemicals	Rubber&plastics	Non-metal	Metal	Machinery	Furniture	Others
1996	35	15	82	42	14	75	12
1997	41	23	89	47	13	108	11
1998	41	26	80	45	17	114	12
1999	38	30	78	45	17	115	10
2000	39	27	76	57	13	113	10
2001	36	27	81	60	7	112	7
2002	41	37	96	72	7	146	7
2003	45	39	111	83	9	157	8
2004	45	42	117	85	9	161	8
2005	51	47	66	86	6	63	9
2006	52	63	137	112	9	183	12
Growth 96-06	49%	320%	67%	167%	-36%	144%	0%