European University Institute

ROBERT SCHUMAN CENTRE FOR ADVANCED STUDIES

EUI WORKING PAPERS

RSCAS 2014/76
Robert Schuman Centre for Advanced Studies
Global Governance Programme-119

Spillovers from agglomerations and inward FDI. A Multilevel Analysis on SSA domestic firms

Marco Sanfilippo and Adnan Seric

European University Institute
Robert Schuman Centre for Advanced Studies
Global Governance Programme

Spillovers from agglomerations and inward FDI. A Multilevel Analysis on SSA domestic firms

Marco Sanfilippo and Adnan Seric

EUI Working Paper **RSCAS** 2014/76

ISSN 1028-3625

Printed in Italy, July 2014
European University Institute
Badia Fiesolana
I – 50014 San Domenico di Fiesole (FI)
Italy
www.eui.eu/RSCAS/Publications/
www.eui.eu
cadmus.eui.eu

Robert Schuman Centre for Advanced Studies

The Robert Schuman Centre for Advanced Studies (RSCAS), created in 1992 and directed by Brigid Laffan since September 2013, aims to develop inter-disciplinary and comparative research and to promote work on the major issues facing the process of integration and European society.

The Centre is home to a large post-doctoral programme and hosts major research programmes and projects, and a range of working groups and *ad hoc* initiatives. The research agenda is organised around a set of core themes and is continuously evolving, reflecting the changing agenda of European integration and the expanding membership of the European Union.

Details of the research of the Centre can be found on:
http://www.eui.eu/RSCAS/Research/

Research publications take the form of Working Papers, Policy Papers, Distinguished Lectures and books. Most of these are also available on the RSCAS website:
http://www.eui.eu/RSCAS/Publications/

The Global Governance Programme at the EUI

The Global Governance Programme (GGP) is research turned into action. It provides a European setting to conduct research at the highest level and promote synergies between the worlds of research and policy-making, to generate ideas and identify creative and innovative solutions to global challenges.

The GGP comprises three core dimensions: research, policy and training. Diverse global governance issues are investigated in *research* strands and projects coordinated by senior scholars, both from the EUI and from other internationally recognized top institutions. The *policy* dimension is developed throughout the programme, but is highlighted in the GGP High-Level Policy Seminars, which bring together policy-makers and academics at the highest level to discuss issues of current global importance.The Academy of Global Governance (AGG) is a unique executive *training* programme where theory and "real world" experience meet. Young executives, policy makers, diplomats, officials, private sector professionals and junior academics, have the opportunity to meet, share views and debate with leading academics, top-level officials, heads of international organisations and senior executives, on topical issues relating to governance.

For more information:
http://globalgovernanceprogramme.eui.eu

Abstract

This paper adopts multilevel analysis to analyse the agglomeration-performance nexus for domestic firms in Sub-Saharan Africa. We show that contextual factors such as country, city and industry together explain up to 30% of the variance in firms' productivity. Our results show also that African firms can take advantage from agglomeration externalities when they locate in cities more densely populated by firms specialized in different sectors (urbanization economies), while their performance worsen when they face direct competition from firms in the same industry. These effects are similar in the services and the manufacturing industries, even if in the latter positive spillovers are found to be conditional to the presence of backward and foreign linkages with nearby firms. Finally, we are also able to show that these effects are magnified when domestic firms locate close to foreign multinationals, especially those coming from the South.

Keywords

Agglomeration economies; Firms' heterogeneity; Sub-Saharan Africa

JEL Classification: D22; F23; O14.

Introduction[*]

Economic activities tend to be unevenly concentrated across space, this being an underlying feature in the process of economic development of nations (World Bank, 2009; Fujita and Thisse, 2013). Compared to advanced economies, in fact, developing countries share some features, such as unbalanced government spending, variance in the economic structure and, generally, high transaction costs, which make geographic polarization more likely to occur (Farole, 2013).

In some circumstances, clustering of economic activities can be viewed as a source of competitive advantage for economic agents, including in particular firms. Since the seminal contribution by Alfred Marshall (1920), a large strand of literature has demonstrated the advantages of clustering to local firms due to the existence of different forms of externalities, such as localization and urbanization economies. More recently, attention is being given to the role of large cities as engines of economic growth and as sources of external economies due to their productivity advantages (Combes et al., 2012; Gill and Goh, 2010), factors that go together with the rising interest in urbanization in developing countries (World Bank, 2009).

This paper tests the existence of such external economies by analysing the nexus agglomeration-productivity for a large sample of domestic firms based in Sub-Saharan Africa (SSA). Our contribution to the existing literature goes in different directions.

First of all, this paper brings new evidence on the relation between agglomeration economies and domestic firms' performance in the context of less developed countries, which so far have received little attention due to the lack of data. Taking advantage of a new and rich database produced by UNIDO, the African Investor Survey, we can extend the analysis of the agglomeration effects on a rich sample of domestic firms from 19 SSA countries (UNIDO, 2012).

Second, the richness of the data allows us to expand on the existing literature, which has focussed on the performance of manufacturing firms within agglomerations, by including also firms in services sectors, which are very likely to benefit from the spatial concentration of economic activities.

Third, we are also able to distinguish an additional source of agglomeration economies, i.e. the co-location with Multinational Enterprises (MNEs), and therefore to test empirically the extent of their spillover effects on domestic firms. This contributes to the large strand of literature on the externalities from FDI in developing countries (Gorg and Greenaway, 2004). We are able to analyse potential differences related to the origin of the foreign investors, i.e. by distinguishing whether there is any difference in agglomerating with Northern or Southern MNEs.

Finally, our contribution to the study of the nexus between agglomeration economies and productivity is also methodological. Following recent developments in urban economics and organizational studies (van Oort et al., 2012; Alcacer et al., 2013), we adopt multilevel analysis to better account for firm heterogeneity and to model the influence of the context in which firms operate on their performance.

Our findings provide empirical support for two major hypotheses. The first is that the context in which firms are embedded, including their geographic area and the sectorial specialization, has an important influence on a firm's performance. We show that, taken together, such contextual factors can explain about 30% of the variance in productivity, more than half of which depends on the geographic location. The second hypothesis is that agglomeration economies result in a mix of positive and negative externalities to African domestic firms. We find robust evidence that firms

[*] We would like to thank Mozghan Raeisian Parvari for her excellent research assistance. We also thank Bernard Hoekman, Matti Sarvimäki and participants to a seminar at UNIDO, May 2014, for their helpful suggestions.

benefit from agglomerations when they locate in urban areas more densely populated by firms specialized in different sectors (urbanization economies), especially with a large variety of economic activities. Conversely, we provide evidence of a competition effect pushing down productivity for firms located close to their main competitors, especially in more concentrated industries. Such results are confirmed when the analysis is run on firms belonging to the services and to the manufacturing sectors separately. In the case of manufacturing firms, however, we find that positive spillovers from agglomeration are more likely to materialize in presence of linkages (backward and forward) with other firms. We confirm therefore the view that linkages act as mediating factors, in line with what has been recently suggested in the literature (Morrisey, 2012).

Our results also suggest that agglomeration with foreign companies generally translates into a stronger impact on domestic firms' performance. Interestingly, the effects are magnified in the case of agglomeration with Southern MNEs. On the one hand, domestic firms seem to take advantage from locating closer to Southern MNEs in diverse sectors, this possibly resulting from a smaller "technology gap", which can give rise to more effective technological spillovers compared to Northern MNEs (Amighini and Sanfilippo, 2014). On the other hand, given that Southern MNEs are likely to provide goods and services that are more accessible to other developing countries (Lipsey and Sjoholm, 2011), this contributes to put them in direct competition with domestic firms producing in the same sector, resulting in above-average negative spillovers.

The rest of the paper is organized as follows. Section 2 reviews the literature on agglomeration economies and their impact on local firms. Section 3 introduces the data, together with some descriptive statistics, and presents the methodology adopted, based on the multilevel analysis. Section 4 discusses the results and Section 5 concludes, drawing some policy implications.

2. Agglomerations and spillovers

The process through which firms tend to concentrate geographically, giving rise to agglomeration economies and externalities, has received substantial attention in the economic literature given its potential implications for local development, industrial policies and firms' performance (Fujita and Thisse, 2013).

The early contribution by Alfred Marshall (1920) emphasized that concentrations of firms in a similar industry give rise to localization economies[1] that are likely to accrue due to the reduced costs of transportation or to the availability of a pool of specialized workers, buyers and suppliers, in turn translating into lower transaction costs. In addition, proximity is likely to foster the circulation of knowledge spillovers, as also suggested by Romer (1990), including new ideas, technologies and business practices. But localization economies are clearly a close proxy for competition, and therefore their overall impact on a firm's performance is ambiguous (Henderson, 2003) since it depends on the nature of the sector and on the degree of market concentration (Porter, 1990).

Localization economies as originally described by Marshall are external to the firms, but internal to the industry in which they operate, and can be related to specialization. Conversely, urbanization economies are independent of the industry and affect all firms located within a given geographic area, based on the principle that it is the diversity of the industries and the actors that stimulates the circulation of knowledge across firms (Jacobs, 1969). In addition to the agglomeration economies described by Marshall, the location of firms within more diversified and dense areas creates greater scope for interacting with diverse actors, such as customers, knowledge-intensive services or related

[1] A substantive evidence has been produced in support of this view, showing that collective efficiency and better performance are more likely to be achieved when such agglomerations are organized in some structured form involving cooperation among sectorally specialized firms and related institutions, such as in the Italian experience with industrial districts (Becattini, 1990).

institutions (Anderson and Loof, 2009). Spatial agglomerations are also more likely to result in pecuniary externalities, for instance by sharing the costs of infrastructure or through forward and backward linkages (Fafchamps and Hamine, 2004).

Looking at the determinants of co-agglomeration choices of US manufacturers, Ellison et al. (2010) have shown that they are most likely to be motivated by the reduced costs of dealing with customers and clients, followed by the opportunity to match with a pool of workers with common skills, and, lastly, by the transfer of ideas. Based on a revised measure of co-agglomeration and on data on Vietnamese firms, Howard et al. (2012) show that technological and knowledge spillovers have a stronger role in a developing country context, especially in high-tech industries. In addition, both studies show also that the relevance of natural advantages of locations, driven by the comparative advantage of access to inputs, cannot be ignored in explaining agglomeration forces.

Whether localization or urbanization economies prevail has an important implication on the developmental potential of the geographic location where they are based, in most cases corresponding to urban areas. Based on an extensive review of the existing evidence, Gill and Goh (2010) conclude that medium-sized cities help domestic firms to exploit localization economies, and this applies mainly to manufacturing industries, while urbanization economies are more likely to be found in larger cities, and in turn foster the proliferation of services. More generally, it has been found that industries based on standardized activities are more likely to take advantage from localization economies and tend to be located in smaller areas compared to productions at the beginning of their life-cycle, who take advantage of larger agglomerations (Henderson, 2003). Externalities from agglomerations often activate a cumulative process, making certain locations more attractive because of their higher productivity. This can happen as a consequence of two main mechanisms. The first is firms' selection, due to higher market competition, which pushes less productive firms out of the market. The second is the agglomeration advantage, which allows surviving firms to enjoy a productivity advantage from co-location. Recent work by Combes et al. (2012) based on French firms demonstrates that – due to firm heterogeneity – the agglomeration advantages distribute unevenly across firms, with those more productive being able to reap stronger benefits.

An additional source of externalities from agglomeration comes from the location choice of foreign Multinational Enterprises (MNEs). There is a large amount of research emphasizing the potential spillovers of FDI through a range of different channels including the creation of forward and backward linkages; the existence of competitive and demonstration effects; the possibility for domestic firms to hire more experienced and skilled workforce; and, more generally, through the transfer of (pecuniary and non-pecuniary) externalities to local firms (Gorg and Greenaway, 2004). Foreign companies bring in advanced production technology and management capabilities, which are potential sources of technological spillovers (Crespo and Fontoura, 2007; Narula and Driffield, 2012). It has been shown that spillover effects from MNEs, either intra- or inter-industry, are more likely to materialize when firms are geographically closer (Farole and Winkler, 2013).

3. Empirical Analysis, data and methodology

3.1 Empirical Framework:

At a more general level, we are interested in understanding the determinants of firm's productivity in the context of African economies. Low levels of productivity represent a binding constraint to the growth potential of firms in the region. Causes of such low productivity range from the poor business environment, to low access to credit or to the structural characteristics of the firms, usually smaller in size and with limited international exposure (Clarke, 2012; Iacovone et al., 2014).

Recent attempts to estimate the determinants of firms' performance in Africa have mainly focussed on their heterogeneous characteristics (Van Biesebroek, 2007), internationalization practices (Clarke, 2012; Foster-McGregor et al., 2013) or on the extent to which they are able to exploit local linkages (Gorg and Seric, 2013). In line with these studies and with the literature on heterogeneous firms (Melitz, 2003), our benchmark model is based on the following general functional relation linking a firm's performance (Y) to agglomeration forces (N), after controlling for a vector of firm-specific factors (Z):

$$Y_{i,j} = f\left(N_{j,c}; Z_i\right) \tag{1}$$

Where i represents the domestic firm, j its industry and c the city.[2] We will base our analysis on an absolute measure of agglomeration, which can be generalized as follows:

$$N_{x,c}^{o} = \sum_x n \tag{2}$$

Where the subscript x represents the generic industry of the agglomeration, while the superscript o refers to the origin of the other firms in the agglomeration (domestic or foreign). This choice is motivated by previous literature, which has adopted the total number of firms in the same region (Siba et al., 2012; Chhair and Newman, 2014) and in the same sector (Henderson, 2003; Fafchamps and Hamine, 2004) as proxies of externalities and competition, respectively. Looking at the number of firms, rather than at the total number of employees, seems a coherent decision in the context of SSA countries, characterized by a large number of SMEs and by a prevalence of unskilled labour. In addition, as externalities are unobservable, it can be argued that it is the firm whose strategic decisions give rise to some kind of spillover, that proxy them at best (Henderson, 2003; Siba et al., 2012)[3]. In the remaining of the analysis we will nonetheless adopt alternative measures, taking into account the number of employees as well, to correct for the differences in firms' size.

Based on the conceptual discussion in section 2, we will test the effects of being part of an agglomeration on firms' performance exploiting the boundaries of *N*. In doing this, we will focus on the two main mechanisms through which agglomeration impacts on firms' performance. The first is competition, which is likely to be significant for firms producing similar types of goods ($x=j$). Competition has an ambiguous impact on firms' performance. On the one hand, competitive pressure in the same market can push firms to organise production more efficiently in order to compete (Porter, 1990). On the other, competition can give rise to negative externalities, potentially leading to reduction in margins and exit from the market. The second effect is the "pure" spillover. Spillovers may assume a variety of forms, including knowledge, technology, workers, and are likely to materialize either when the firms operate in similar ($x=j$) of different sectors ($x\neq j$).

We expect spillover effects to have a stronger impact in our sample given that some studies have highlighted that there is a larger potential in developing countries, where firms operate away from the technology frontier (Siba et al., 2012). However, as recently argued by Fafchamps and Soderbom (2014) in a study on the diffusion of business practices among African firms, we should also be aware that geographic proximity per se does not automatically translate into greater spillovers, and that other factors should be taken into account in the analysis.

Table 1 provides a description of the different measures of agglomeration that, based on (2), will be tested in the empirical analysis, detailing the expected impact on firms' performance.

---

2 We refer to the city as our geographic unit, and not to the district or other smaller units, because of the lack of information on the full address for a large number of firms in the survey. In addition, for a number of the remaning cases, even in presence of the full address, we were not able to correctly geocode firms, due to a scarce coverage of existing specialized softwares for remote areas in Africa.

3 An additional concern of using the total number of workers as a proxy for agglomeration is endogeneity, as an increase in productivity might induce firms to expand and hire more workers.

Table 1. Main measures of agglomeration

Measure	Description	Expected effect
N_c	Total n. of firms located in the same city*	+
$N_{j,c}$	Total n. of firms in the same city and in the same industry*	+/-
$N_{p,c}$	Total n. of firms in the same city and producing the same product (*),(**)	+/-
$N_{x,c}$	Total n. of firms in the same city and in a different industry	+
$N_{j,c}^{f}$	Total n. of MNEs in the same city and in the same industry	+/-
$N_{x,c}^{f}$	Total n. of MNEs in the same city and in a different industry	+

*The total number does not include the firm itself

**This measure at the product level can be only computed for manufacturing firms.

Most of the empirical work on the impact of agglomerations on firm performance focuses on the learning mechanism and emphasizes the role of spillovers. Existing evidence, mostly on developed countries, tend to support the view that agglomerations work as a shifter in firm's production function, this being in most cases independent on the sector (Henderson, 2003; Anderson and Loof, 2009). Evidence from developing countries is nonetheless becoming more consistent, and is generally reporting similar results (Farole, 2013). Due to poor availability of data, work on agglomerations of firms in Africa has been limited to anecdotic evidence or few case studies, initially focussing on the role of industrial clusters (McCormick, 1999; Yoshimo, 2012)[4], but is now growing thanks to greater availability of firm-level information.

Relevant benchmarks for this study are some recent works looking at agglomeration externalities in the forms of competition and spillover on the productivity of developing country firms. Two works on firms from Cambodia (Chhair and Newman, 2014) and Vietnam (Howard et al., 2014), for instance, tend to support the view that firms within clusters enjoy a premium in terms of productivity *vis à vis* non-clustered firms, but also that this happens through a variety of different mechanisms and tend to be stronger for some firms. Fafchamps and Hamine (2004) analyse firms from Morocco and find empirical support for the returns from localization externalities hypothesis, results being robust to the adoption of different measures of agglomeration. However, they show also that the net impact of competition can be negative. Siba et al. (2012) analysis based on Ethiopian firms shows instead that the competitive effects from the agglomeration of specialized producers translates into an increase of productivity, but a reduction of prices. In line with the previous study, they show also that the competitive effects of agglomerations on profit margins might overcome the advantages found in terms of technical efficiency, concluding that firms might not be well motivated to join clusters endogenously.

[4] The most comprehensive work so far is a recent report by the World Bank (Yoshimo, 2012), which is based on the analysis of five country case studies and seems to find support for the existence of Marshallian economies. The analysis finds a positive correlation between location within selected industry-specific clusters and a range of indicators of firms' performance, pointing this advantage to be a consequence of a better accumulation of capital within the boundaries of the clusters.

3.2 Empirical specification: the Multilevel approach

Based on the discussion in the previous section, we derive our baseline empirical specification:

$$Y_{i,j} = \beta_1 N^o_{x,c} + \beta_2 Size_i + \beta_3 Age_i + \beta_4 Fam_i + \beta_5 R\&D_i + \beta_6 Exp_i + \beta_7 Skill_i + \gamma_k + \delta_j + \varepsilon_{i,j} \tag{3}$$

Where our dependent variable is labour productivity, measured as the ratio of total sales over the number of employees. Using a revenue-based measure of productivity raises some important issues, as it captures both differences in productivity and in mark-ups across firms.[5]

The vector of variables Z in equation (1) includes a number of controls to account for firms' heterogeneity. They include standard variables such as the age (*Age*) and the size (*Size*) of the firm; both expected to be positively correlated with productivity (Melitz, 2003), and the family ownership (*Fam*), which usually has a negative impact on firms' performance. In addition, we account for the innovation effort (*R&D*) and the internationalization status (*Exp*), both of which have been previously identified as significant predictors of performance, including in the African context (Gorg and Seric, 2013; Foster-McGregor et al., 2013). We also test whether the skill intensity of workers in a firm influences its level of productivity (*Skill*). Finally, the specification includes both country (γ_k) and sector (δ_j) fixed effects. Variable description and their summary statistics are reported in Table A1 in the appendix.

A well-known issue when estimating (3) by means of a standard OLS is potential endogeneity. This is due to the self-selection of more productive firms into larger and better performing agglomerations. An additional (and related) issue, still partially unexplored in the literature on the agglomeration-performance nexus, is that most analyses do not properly take into account the typically complex structure of the data, which tries to observe the effect of the context (the macro dimension) in which firms operate on the individual firms' performance (the micro dimension).

Recent advances in urban economics and organizational studies have documented the advantages of adopting multilevel analysis to best take into account firms' heterogeneity and to model the influence of the context on firms' performance (van Oort et al., 2012; Alcacer et al., 2013).

Multilevel analysis models the micro and the macro dimensions of the data simultaneously. While clustering the error term assumes homogeneous correlation structures for all the groups and fixed effects estimators allows for unique variability within groups, a multilevel approach controls for the larger complexity given by the hierarchical structure in the data. This, in turn, translates in the adoption of a maximum likelihood estimator leading to more efficient estimates of the coefficients and their standard errors (Snijders and Bosker, 1999; Maas and Hox, 2004). Looking specifically at the features of our data, the adoption of Multilvel analysis has a number of advantages. The first is that it is possible to model the impact of the context on outcomes. In our specific case, this allows to understand why firms within agglomerations are more likely to perform similarly. It will also allow to identify how much of the variance in firms' productivity can be explained by between-firm or between-agglomeration variance. Second, they account more properly for unobserved heterogeneity thanks to the inclusion of random (together with fixed) coefficients (Alcacer et al., 2013). In addition, van Oort et al. (2012) show that multilevel analysis can contribute to solve the "agglomeration-performance ambiguity" in that it controls for the circular process that link firm and location's performance.

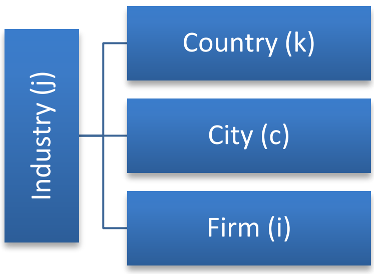
Methodologically, the rationale for adopting a multilevel model is due to the unrealistic assumption that in a generic form like (3) the deviation $\varepsilon_{i,j}$ from $Y_{i,j}$ is uncorrelated within subjects (Rabe-Hesketh and Skrondal, 2012). Multilevel models address such dependency by splitting the residual into two

[5] The drawback of using such aggregate measures is that they do not allow to determine if the competitive effect from agglomeration hits the technical efficiency or pushes prices downward.

uncorrelated components, including a permanent component ζ_i, which measures the random deviation of subject i's mean from β, and an idiosyncratic component, which is specific to each subject across all the dimensions j. But multilevel models can accommodate more complex structures too, including with more than two levels (by adding up a random coefficient for each level) as well as cases in which subjects are nested in non-hierarchical structures, i.e. they are cross-classified by two or more factors (Rabe-Hesketh and Skrondal, 2012).

In light of this, and given the peculiar nature of our data, we model them according to a structure in which the firms are nested into a hierarchical structure including at the top the country of origin and the city, both likely to affect the performance, especially when they correspond with the spatial boundaries of agglomerations (Figure 1). We also add an additional non-hierarchical dimension, the industry, which is not nested in the structure, but across it, given that the measures of agglomerations are sector specific and that firms belonging to a given industry can be nested within the same country-city.

Figure 1. Structure of the data



Our final specification is therefore the following:

$$Y_{i,j,c,k} = \beta_1 N^o_{x,c} + \beta_2 Size_i + \beta_3 Age_i + \beta_4 Fam_i + \beta_5 R\&D_i + \beta_6 Exp_i + \beta_7 Skill_i + \delta_j + \zeta_{i,j(c,k)} + \zeta_{i,c} + \zeta_{i,k} + \epsilon_{i,j,c,k} \quad (4)$$

Where the first part of the equation reports the fixed part of the model with the predictors. We still include industry dummies (δ_j) in the model. The second part of the equation reports the random coefficients, both *iid* distributed with mean zero and constant variance.

An important feature of multilevel models is that – even in the absence of explanatory variables at higher levels of aggregation– they still perform better than a standard model, which violates the assumption of independence of all observations when data are nested (Rabe-Hesketh and Skrondal, 2012). It is possible to estimate the dependence by specifying an empty model from which the variances of the lower and of the higher level error terms are then retrieved. The Variance Partition Coefficient (VPC) is then computed according to the following formula (for example for the city):

$$VPC = \frac{\sigma^2_{i,c}}{(\sigma^2_{i,j(c,k)} + \sigma^2_{i,c} + \sigma^2_{i,k})} \quad (5)$$

Where the numerator includes the level-specific variance, and the denominator is the total one. Given the above, table 2 reports the results obtained by running an empty model to retrieve the VPCs for the four levels considered. These results show that – other things equal – the heterogeneity in characteristics of the firms explain the largest part (around 69%) of their performance. Strikingly, such results tell also that there is a large share of the variance of firms' productivity that is affected by the context in which they operate, calling for a stronger focus on their behaviour. Between industries variance accounts for 14%, while the location of the firm contributes together (accounting for both the country and the city) to explain 17.4% of the variation in firms' performance.

Table 2. VPCs at the different levels and their contribution to total variance

Between firm variance	68.7%
Between industry variance	13.9%
Between city variance	6.7%
Between country variance	10.7%
Total	100%

3.3 Data and descriptive statistics

We use original firm-level data collected through the UNIDO Africa Investor Survey 2010 across 19 Sub-Saharan Africa countries[6]. We use both the Foreign- and Domestic Investor Survey data, which contain a rich set of information on a large sample of foreign and domestic owned firms. The collection of the dataset followed a rigorous survey methodology in terms of stratified sampling (on three dimensions: sector, size and ownership) and interview techniques (face-to-face interviews with top-level managers of foreign- and domestically-owned firms). The sample was constructed in order to be representative of public and private for profit firms with 10 or more employees[7].

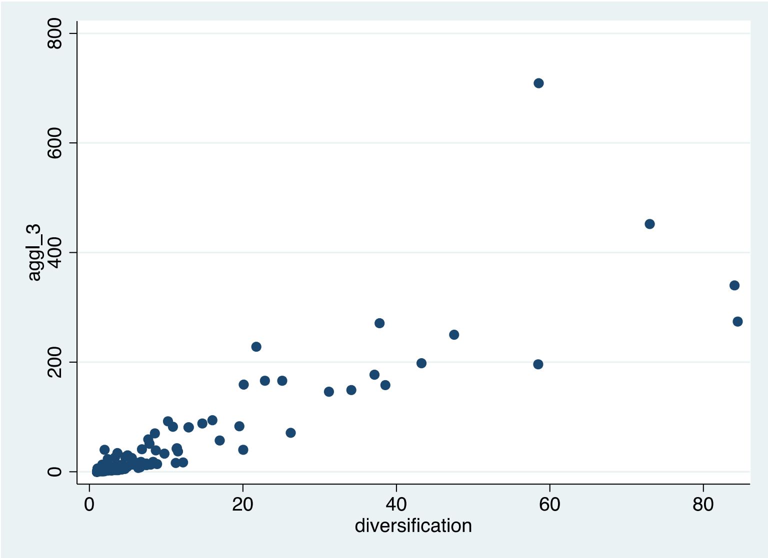
The Africa Investor Survey provides specific information on agglomeration and inter-firm linkages in the host country. Specifically, the survey features detailed information on firms' location, industry and product classification as well forward and backward linkages, among others.
There is one disadvantage, however. Currently, the data are only available for a cross section in 2010. Hence, while we can use the data to unearth and describe some hitherto unknown relationships, we are careful to avoid interpreting these as causal effects. Nevertheless, we feel that the relationships are sufficiently interesting and, importantly, policy relevant to justify our analysis.

Moving to the construction of the agglomeration variables, some descriptive statistics show that the cities with the larger concentration of firms are the main capital cities in the countries, including in particular Kampala, Nairobi, Addis Ababa, Dakar (Figure A1 in the appendix). When it comes to the presence of foreign companies, Kampala keeps its leadership followed by Accra. It is interesting to observe that larger agglomerations are also generally characterized by a large variety of economic activities. Indeed, if we plot together the size of agglomerations with an index of diversification of the economic activities performed in the city, we show a straight relation between the two variables (Figure 2), even if there is some notable distinction, such as Accra, Lagos or Maputo, all hosting a large number of firms, but with a smaller number of industrial activities.

6 Burkina Faso, Burundi, Cameroon, Cape Verde, Ethiopia, Ghana, Kenya, Lesotho, Madagascar, Malawi, Mali, Mozambique, Niger, Nigeria, Rwanda, Senegal, Tanzania, Uganda, Zambia.

7 An oversampling of relatively large firms (> 100 employees) has been adopted.

Figure 2. Agglomeration size and the diversification of economic activities*

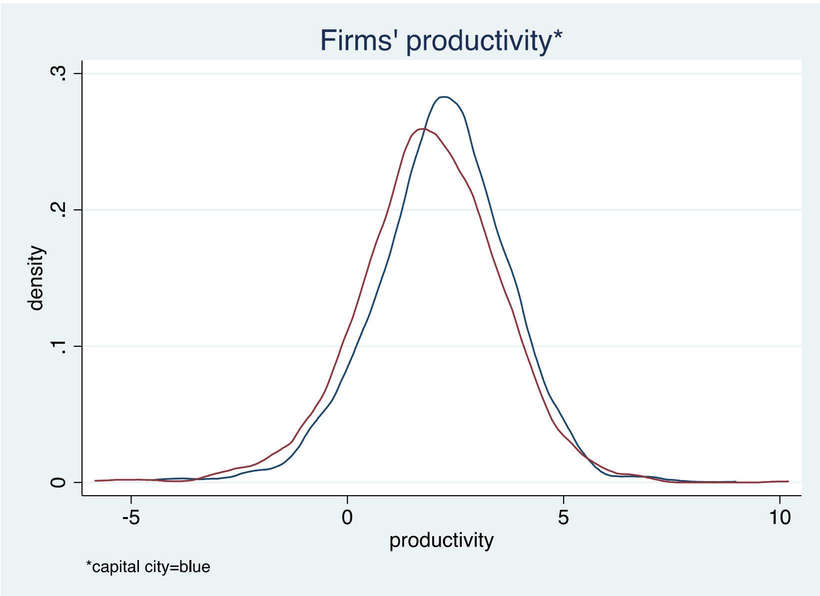


Source: Authors' elaboration

* See section 4.1 for a description on how the diversification index is constructed.

Finally, it is useful for the remainder of our analysis to consider whether firms in larger cities enjoy some kind of productivity premium, as suggested by existing evidence (Combes et al., 2012). In figure 3 we plot the density distribution of firm's productivity, simply distinguishing if they are located in capital cities or elsewhere. In our sample of SSA countries, capital cities are in most cases the most populated in the country, and the ones hosting the larger agglomerations (see also figure A1). The graph shows that the distribution of firms in the capital cities is shifted rightward and indicates that more productive firms are indeed located in larger cities.

Figure 3. Distribution of domestic firms' productivity by location



Source: Authors' elaboration

4. Results

4.1 Benchmark specification

Table 3 reports results of our main specifications on the impact of agglomerations on productivity. In the first two columns, results have been obtained through a standard OLS with robust standard errors. The following columns report instead results based on the multilevel model, on which we will base our comments. Overall, both sign, magnitude and significance of the coefficients are quite robust to the different methodologies adopted, but there is an improvement in the standard errors when moving to Multilevel.

Table 3. Main Results

	(1) OLS	(2) OLS	(3) Multilevel	(4) Multilevel	(5) Multilevel	(6) Multilevel
Size	0.181***	0.182***	0.172***	0.173***	0.172***	0.174***
	(0.034)	(0.034)	(0.046)	(0.046)	(0.046)	(0.046)
Age	0.139***	0.134***	0.142***	0.141***	0.142***	0.143***
	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)
Exp	0.501***	0.496***	0.482***	0.482***	0.484***	0.480***
	(0.072)	(0.072)	(0.071)	(0.071)	(0.071)	(0.071)
Family	-0.232***	-0.227***	-0.231***	-0.230***	-0.228***	-0.230***
	(0.053)	(0.053)	(0.074)	(0.073)	(0.073)	(0.074)
skill	0.007***	0.007***	0.007***	0.007***	0.007***	0.007***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
R&D	0.031	0.037	0.029	0.031	0.028	0.023
	(0.064)	(0.064)	(0.065)	(0.064)	(0.065)	(0.064)
N_c	0.000**		0.001***			
	(0.000)		(0.000)			
$N_{j,c}$		-0.007**		-0.004		
		(0.003)		(0.004)		
$N_{x,c}$		0.001***		0.001**		
		(0.000)		(0.000)		
Div_c					0.076**	
					(0.035)	
$C_{i,j}$						-0.279***
						(0.099)
Constant	1.810***	1.788***	1.048***	1.043***	0.932***	1.271***
	(0.248)	(0.247)	(0.233)	(0.229)	(0.224)	(0.258)
Observations	3,281	3,281	3,281	3,281	3,281	3,281
R-squared	0.274	0.276				
Country effects	Yes	Yes	Yes	Yes	Yes	Yes
Industry effects	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

As far as the control variables are concerned, results are pretty much in line with a priori expectations. Larger firms are more productive than smaller ones, and the same is true for firms with longer experience. Another stylized fact from the literature, i.e. that family owned firms experience low levels of productivity, is strongly supported by our data. Still, we confirm that the nexus internationalization-productivity holds true also in the context of SSA countries, showing that being an exporter guarantees a productivity premium to the firm. This is not surprising, as previous research using the same data has found similar results, and indeed the magnitude of the coefficient is in line with such work (Foster-McGregor et al., 2013). Also the coefficient measuring the skill ratio, a proxy for human capital endowments, reports the expected sign, again in line with previous research. Lastly, we don't find robust results on the nexus between productivity and innovation, contrary to existing literature pointing to a consistent positive relation between the two variables (Grossman and Helpman, 1991). The coefficient of the variable representing the R&D effort, though positive, is not significant in any specification. This is consistent with what has been found in other research using the same data (Gorg and Seric, 2013).

Moving now to our variables of interest, results seem to support the view that the context explains differences in productivity between firms (see Table 3) and that being located into a large agglomeration of firms, independently on their sector, can translate into positive externalities. As indicated by the description of the data, larger agglomerations are most likely to be settled in larger and more productive cities, which in our sample often correspond to the administrative capital. In light of this, we can try to read this first result as the confirmation of the more general view that firms benefit from the proximity to large markets (Krugman, 1991). This is also consistent with the evidence that firms and workers based in larger cities enjoy a productivity premium (Combes et al., 2012). This first result is therefore not surprising, and clearly deserves a deeper investigation.

More interesting results come when we distinguish between competitive and spillover effects (column 4). Specifically, we find quite robust evidence that urbanization economies, i.e. the agglomeration with other firms belonging to different sectors, are responsible for the positive impact found in column (3). In order to understand whether this effect can be attributed to the scale of the agglomeration and/or if there is also a composition effect, we try to account for sectorial variety, constructing a diversification index, defined as:

$$Div_c = 1/C_c \qquad (6)$$

Where C is a Herfindal index ranging from 0 (maximum diversity) to 1 (maximum concentration), which is computed taking into account the number of employees at the level of the sector ($L_{j,c}$) and the city (L_c):

$$C_c = \sum_j \left(\frac{L_{j,c}}{L_c}\right)^2 \qquad (7)$$

As expected, the index has a positive and significant coefficient[8], indicating that the larger the variety of economic activities performed within the boundaries of the city, the higher is the productivity of local firms.

Conversely, at a first sight, we find no significant evidence on localization economies and the competition effect, measured by the total number of firms in the same industry.[9] The sign of the latter coefficient is negative, which may reflect that in unsophisticated markets like in SSA countries, the competition effect works mostly in the direction of reducing margins rather than driving firms towards higher efficiency. As suggested by Combes et al. (2012), this can be even more the case in larger cities, where competition is tough and less productive firms are more likely to exit the market while larger firms tend to become more productive.

In order to understand better the nature of the competitive effect, we estimate an additional specification that includes a measure of industry concentration:

---

[8] The diversification index reported in the table of results has been transformed in log.

[9] This finding is somewhat in contrast with previous findings on Ethiopian firms (Siba et al., 2012). Besides the differences in the samples analysed, one reason is that the authors are able to distinguish between a positive impact due to competition on production efficiency and a negative one, on prices.

$$C_{i,j} = \sum_{i,j} (\frac{L_{i,j}}{L_j})^2 \quad (8)$$

Where $C_{i,j}$ is measured as a Herfindal index, taking the value of 0 if all the firms in the sector share the same size (in terms of number of employees) and 1 if all of them are concentrated in just one firm. Contrary to the Schumpeterian view that industries with a monopolistic competition and market concentration are those with higher externalities, we show that, still, positive effects arise from competition, as in Siba et al. (2012). More specifically, we are able to add that this happens only in those industries where markets are closer to perfect competition, which in our sample include for instance food processing industry; publishing; construction and retail trade.

4.2 Does it matter where your neighbour comes from?

After having controlled for the overall impact of agglomeration according to the two main mechanisms analysed, in what follows we take advantage of the richness of our data and try to distinguish whether the competitive and the spillover effects discussed in the previous section are likely to be influenced by the presence of foreign MNEs. In less developed countries, agglomerations provide some incentives for MNEs to invest and can substitute for inefficient policies (Yehoue, 2005). In turn, as discussed in section 2, the externalities stemming from co-location with MNEs can be significantly higher, especially for firms far from the frontier, even if this is highly dependent on several factors including the motivations of the investors (Aitken and Harrison, 1999) or the absorptive capacities of domestic firms (Morrisey, 2012). A recent empirical analysis on SSA firms seems to support this view, since it shows that firms within agglomerations are more likely to maximize the spillover potential from FDI (Farole and Winkler, 2013).

In what follows, therefore, we will report the estimation results of (4) distinguishing whether the sources of competition and spillovers arise from agglomerations made up by other domestic firms ($N^d_{x,c}$) or from foreign-owned companies ($N^f_{x,c}$).

Table 4. Results disaggregated by domestic and foreign ownership

	(1)	(2)	(3)	(4)	(5)
Size	0.173***	0.172***	0.171***	0.172***	0.172***
	(0.046)	(0.045)	(0.045)	(0.046)	(0.045)
Age	0.141***	0.142***	0.140***	0.140***	0.142***
	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)
Exp	0.481***	0.484***	0.481***	0.479***	0.485***
	(0.070)	(0.073)	(0.074)	(0.073)	(0.073)
Family	-0.229***	-0.232***	-0.231***	-0.234***	-0.232***
	(0.073)	(0.074)	(0.073)	(0.074)	(0.074)
skill	0.007***	0.007***	0.007***	0.007***	0.007***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
R&D	0.031	0.030	0.032	0.031	0.027
	(0.064)	(0.065)	(0.064)	(0.065)	(0.065)
$N^{d}_{j,c}$	-0.004		-0.005		
	(0.005)		(0.004)		
$N^{d}_{x,c}$	0.001*		-0.001		
	(0.001)		(0.002)		
$N^{f}_{j,c}$		-0.007	-0.005		
		(0.008)	(0.008)		
$N^{f}_{x,c}$		0.002***	0.003*		
		(0.001)	(0.002)		
$N^{f}_{j,c}$(north)				0.001	
				(0.009)	
$N^{f}_{x,c}$(north)				0.002***	
				(0.001)	
$N^{f}_{j,c}$(south)					-0.024***
					(0.008)
$N^{f}_{x,c}$(south)					0.003***
					(0.001)
Constant	1.045***	1.057***	1.085***	1.074***	1.052***
	(0.229)	(0.235)	(0.226)	(0.238)	(0.237)
Observations	3,281	3,281	3,281	3,281	3,281
Industry dummies	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Since the results of the control variables remain robust, we focus on the subset of variables of interest. Surprisingly, at a first glance, we do not find big differences in the signs and the significance of the coefficients compared to the results of the general model in Table 3. Co-location with other firms involved in different industries is found to be a significant driver of positive spillover, independently on whether the other firms are domestically or foreign owned. Similarly, the coefficient representing the competitive effect of agglomeration continues to be non-significant for both the groups of firms, but keeps suggesting that the likely effect is in any case negative. This said, it is nonetheless useful to observe that the magnitude of the two coefficients is stronger in the case of foreign owned companies. As might be expected, the marginal effect of adding one more unit to the agglomeration affects domestic firms differently according to the origin of the new firm. More specifically, the entry in a city of an additional MNE in a different (or in the same) industry results in an increase (decrease) in productivity of 0.2% (0.7%), while the same effect related to the entrance of an additional domestic firm is 0.1% (0.4%).

This last result does not come at a surprise when looking at the positive spillover, given that the stock of knowledge flows and other skills brought in by foreign firms are realistically meant to be higher in view of their unique set of competitive advantages (Crespo and Fontoura, 2007). On the other hand, it is a bit more surprising to see that – through not significant – the extent of the competition effect looks stronger when compared to agglomerations with other domestic firms. Indeed, following what has just been said, one would have expected that due to the higher technology gap the extent to which domestic firms compete with foreign multinationals should be marginal, at least compared to direct competition with other domestic firms. Also the results of the survey seem to suggest something similar. When domestic firms were asked to report which is the main source of competition, 67.6% of them responded that is another local company, while only a 16.5% has indicated a foreign company (the remaining has mentioned imports).

This said, it should be noted that MNEs investing in SSA countries in recent years have become more heterogeneous. This reflects the rising share of South-South FDI, with MNEs from emerging markets like China, India, South Africa or other east Asian countries entering the continent with a variety of motivations and new approaches. It has been argued that, compared to North-South FDI, South-South FDI potentially brings more positive effects to host economies. This is due, for instance, to the fact that lower institutional distance fosters business integration, as shown by a recent work looking at the determinants of backward linkages between MNEs and domestic firms in SSA (Perez-Vilar and Seric, 2014). In addition, given a smaller "technology gap", Southern MNEs can generate more effective technological spillovers (Amighini and Sanfilippo, 2014). Southern MNEs are likely to provide goods and services that are more accessible to other developing countries (Lipsey and Sjoholm, 2011), and this is perhaps a feature that can put them in direct competition with domestic firms.

In light of the above discussion, we exploit the richness of our database and further disaggregate results of column (2) to check the effect of agglomeration with Northern and Southern MNEs. Results are very interesting, for both the effects studied. First, they seem to support the idea that the potential spillovers to domestic firms are slightly higher if they co-locate with Southern MNEs. If we think of the spillover as the flow of knowledge or technologies, this result perfectly fits with the idea that the impact is higher when the gap between domestic and foreign firms is smaller. On the other hand, possibly due to the same motivation, we find that if the sectorial specialization is the same, domestic firms find in Southern MNEs strong competitors in the local market. This competition potentially translates in a reduction in margins not compensated by an increase in efficiency, leading thus to lower levels of productivity. This effect looks particularly strong: our results suggest that the entry of an additional Southern MNE corresponds to a decrease of productivity by 2.4%.

Conversely, in line with the ex-ante expectations, we do not find evidence of a competitive effect from advanced countries' MNEs. In this case, in fact, the coefficient even changes its sign compared to the general results, even if it keeps being not significant.

4.3 A sector-based analysis

As a final step of our empirical analysis, we are now interested in understanding whether the impact observed in the previous sections can be generalized, or if domestic firms are affected differently according to their main sector. In fact, most of the literature on agglomeration economies has focused so far on the manufacturing sector only, ignoring the role of the services, despite services are even more spatially concentrated than manufacturing[10] (Gill and Goh, 2010).

Table 5 reports the results for firms in the manufacturing sector only. Overall, results are somewhat in line with the previous, but they are statistically weaker, as most of the variables measuring

[10] This, according to Gill and Goh (2010: 246) is due to two main reasons. The first is that service firms need less land per employees. The second is that, by nature, service firms need to locate close to other firms, both producers and other complementary services, which are often among their major customers.

agglomerations lose their significance. This notwithstanding, we are still able to add some relevant insights on this group of firms. The first is that for manufacturing firms we are able to compute a further, and more disaggregated, level of agglomeration that is constructed as the number of firms producing the same product ($N_{p,c}$)[11]. This additional information allows to provide a better specification for the competitive effect, compared to one including all firms operating within the same 2-digit industry. As a matter of fact, when we introduce this new variable in the model, we find evidence of a strong and significant negative effect on productivity. This means that competition among firms specialized in the production of the same product lines is a driver of low performance, suggesting that price competition and the erosion of margins prevail over the expected increase in efficiency when firms find themselves to share a very specific market with a large number of competitors.

[11] This is possible because most of the manufacturing firms have carefully described their main product in response to a specific question of the survey. For each product indicated, we have then attributed a common label, referring to the 6-digit classification of the Harmonised system to make them comparable over firms.

Table 5. Results for the manufacturing sector

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Size	0.344***	0.330***	0.329***	0.344***	0.344***	0.329***	0.342***	0.328***
	(0.052)	(0.059)	(0.058)	(0.052)	(0.052)	(0.049)	(0.051)	(0.049)
Age	0.101***	0.078**	0.078**	0.101***	0.103***	0.119***	0.102***	0.118***
	(0.029)	(0.030)	(0.031)	(0.030)	(0.029)	(0.032)	(0.029)	(0.032)
Exp	0.389***	0.521***	0.518***	0.388***	0.389***	0.340***	0.376***	0.333***
	(0.089)	(0.097)	(0.096)	(0.089)	(0.088)	(0.101)	(0.092)	(0.103)
Family	-0.180*	-0.196**	-0.196**	-0.179*	-0.183**	-0.191**	-0.184*	-0.194**
	(0.093)	(0.079)	(0.078)	(0.093)	(0.093)	(0.084)	(0.094)	(0.085)
skill	0.007***	0.005**	0.005**	0.007***	0.007***	0.008***	0.007***	0.008***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
R&D	-0.005	0.015	0.014	-0.006	-0.007	-0.011	-0.016	-0.011
	(0.066)	(0.078)	(0.078)	(0.066)	(0.066)	(0.077)	(0.066)	(0.076)
$N_{j,c}$	-0.003							
	(0.004)							
$N_{x,c}$	0.001	0.001*	0.001					
	(0.000)	(0.000)	(0.000)					
$N_{p,c}$		-0.023**	-0.023*					
		(0.012)	(0.012)					
$N_{(j-p),c}$			0.004					
			(0.005)					
$N^d_{j,c}$				-0.003				
				(0.005)				
$N^d_{x,c}$				0.001				
				(0.001)				
$N^f_{j,c}$					-0.005			
					(0.008)			
$N^f_{x,c}$					0.001			
					(0.001)			
N^d_c*backward						0.000*		0.000
						(0.000)		(0.000)
N^d_c*foward						-0.000**		-0.000**
						(0.000)		(0.000)
N^f_c*backward							0.000	-0.000
							(0.000)	(0.000)
N^f_c*forward							0.001**	0.000***
							(0.000)	(0.000)
Constant	0.529**	0.647**	0.698***	0.551**	0.562**	0.631***	0.643**	0.642***
	(0.245)	(0.288)	(0.264)	(0.240)	(0.267)	(0.243)	(0.254)	(0.240)
Obs	1,624	1,212	1,212	1,624	1,624	1,414	1,624	1,414
Industry dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

But this is not the only interesting finding for the group of manufacturing firms. Indeed, one can realistically assume that spillovers or other external economies arising from co-location with other companies, especially foreign, are more difficult to materialize in productive activities, given that labour mobility is limited as it is the flow of tacit knowledge (to not talk about the low absorptive capacities). This is especially true in specific the context of our survey, where about two thirds of domestic manufacturers are involved in low-technology activities.

In light of this, and along the lines of Morrisey (2012), we try to look at whether such spillovers are more likely arise in presence of linkages between firms. This is due to the fact that linkages are most frequent in the manufacturing, and can eventually give rise to positive spillovers due to the improvements in standards and production (backward linkages) or to the adoption of more sophisticated inputs (forward linkages) (Morrisey, 2012). And a recent work based on the African Investor Survey has in fact suggested that linkages (with foreign companies) seem to be relevant vehicles of learning and, in fact, they are found to improve manufacturing firms' performance (Gorg and Seric, 2013). In light of the above discussion, we interact our variables representing agglomerations with some dummies, equal to 1 if the domestic manufacturer has direct linkages with at least one other company based in the same area. Thanks to the richness of the data, we are able to disaggregate the extent of linkages, according to whether they are backward or forward, as well as if they involve a domestic or a foreign company.

Results, reported in columns 6-8, are straightforward, since they seem to support the view of linkages as mediating factors to activate spillovers in the manufacturing sector (Morrisey, 2012). More specifically, we find that firms' productivity is enhanced by co-location with other domestic firms in presence of backward linkages. No significant effects are recorded for backward linkages with foreign companies. These results do not come as a surprise if we refer to the original data. Backward linkages with foreign firms are less frequent than with domestic firms (on average, each local producer has 53 domestic buyers and only 4.3 foreign). In addition, domestic firms can specialize in the supply of more standardized productions for foreign companies, while their supply is more strategic to other domestic firms. Again, if we look at the data from the questionnaire, in 62% of the cases the domestic buyers provides support deemed useful to "upgrade the efficiency of (the) production process" against the 17% of cases for foreign buyers.

Conversely, we find that the agglomeration with foreign firms enhances productivity in case of forward linkages, while the opposite happens for domestic firms. These results are quite tricky to be interpreted together. The opposite sign of the coefficient of forward linkages, in particular, seems to suggest that there is an intrinsic difference in the quality of the inputs sourced, and that buying intermediate or final inputs from MNEs taking advantage of their proximity favours the transfer of knowledge and the learning process. The latter result is in line with the findings by Gorg and Seric (2013), who show that sourcing superior inputs from foreign companies allows domestic firms to produce in a more efficient way.

Moving now to the analysis of the service sector, results seem to reflect more closely those reported in tables 3 and 4, except for the lack of significance of the coefficient representing firms' size. We find in particular that service firms tend to benefit from urbanization economies, especially when they locate close to foreign companies. In addition, we find that positive spillovers arise when service firms co-locate with other firms producing knowledge intensive services ($N_{kis,c}$). On the other hand, we show also that domestic service firms are particularly affected by the direct competition of southern multinationals. This displacement effect can be due to an inherent advantage of EMNEs in the provision of services more targeted to the local needs, thanks to cultural and geographical proximity on the one hand, and the exploitation of scale economies on the other (Barnard, 2008).

Table 6. Results for the service sector

	(1)	(2)	(3)	(4)	(5)	(6)
Size	0.016	0.016	0.018	0.014	0.016	0.013
	(0.061)	(0.060)	(0.061)	(0.060)	(0.059)	(0.060)
Age	0.184**	0.185**	0.186**	0.182**	0.188**	0.189**
	(0.078)	(0.077)	(0.078)	(0.077)	(0.077)	(0.076)
Exp	0.676***	0.677***	0.680***	0.673***	0.660***	0.672***
	(0.143)	(0.145)	(0.144)	(0.143)	(0.138)	(0.142)
Family	-0.250***	-0.251***	-0.249***	-0.251***	-0.259***	-0.249***
	(0.086)	(0.086)	(0.086)	(0.086)	(0.086)	(0.086)
skill	0.007***	0.007***	0.007***	0.007***	0.006***	0.007***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
R&D	0.107	0.104	0.106	0.106	0.097	0.097
	(0.150)	(0.151)	(0.148)	(0.151)	(0.150)	(0.152)
$N_{j,c}$	-0.005	-0.005				
	(0.009)	(0.010)				
$N_{x,c}$	0.001*					
	(0.001)					
$N_{kis,c}$		0.002*				
		(0.001)				
$N^d_{j,c}$			0.000			
			(0.015)			
$N^d_{x,c}$			0.001			
			(0.001)			
$N^f_{j,c}$				-0.019		
				(0.014)		
$N^f_{x,c}$				0.002***		
				(0.001)		
$N^f_{j,c}$(north)					-0.011	
					(0.027)	
$N^f_{x,c}$(north)					0.004***	
					(0.001)	
$N^f_{j,c}$(south)						-0.037***
						(0.009)
$N^f_{x,c}$(south)						0.005***
						(0.001)
Constant	0.951***	0.960***	0.955***	0.953***	0.977***	0.983***
	(0.316)	(0.320)	(0.317)	(0.315)	(0.320)	(0.322)
Obs	1,530	1,530	1,530	1,530	1,530	1,530
Industry dummies	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

4.4 Robustness checks

In what follows, we test the stability of our previous results against the adoption of different indicators to measure both firms' performance and agglomeration forces.

As discussed in section 3.1, previous literature has often used the number of employees as a proxy for agglomeration, given that the mobility of workers across firms and the exchange of information between individuals, both formally and informally, could favour the transmission of spillovers (Fafchamps and Hamnie, 2004; Fujita and Thesse, 2013). In light of this, we run our main specifications replacing the set of agglomeration measures using the total number of workers rather than the number of firms (Table A2, columns I-II, in the Appendix). Given the high degree of correlation between the two measures, it is not surprising to discover that the coefficients tend to report the same signs. Thus, we can confirm that the larger the agglomeration the stronger the effect on a firm productivity and, again, especially when other economic activities are not concentrated in the same industry. The only noticeable difference with results in table 3 is that the coefficient of the competition effect (i.e. the number of employees in the same industry) is not statistically significant, this possibly being linked to the fact that competition is mostly determined by the number of competitors, rather than on the overall size of the industry.

Moving to our dependent variable, we try to adopt an alternative measure of firm efficiency. Specifically, we construct a simple estimation of total factor productivity using a constant return to scale Cobb-Douglas production function:

$$TFP_i = Y_i/(K_i^{\alpha} L_i^{1-\alpha}) \tag{7}$$

Where Y, the output, is measured by sales on turnover, L is the total number of employees and K fixed assets, assuming a share of 2/3 for the former and 1/3 for the latter.

Also in this case, running our model with the same set of dependent variables do not affect the results, showing similar correlations between agglomeration forces and firms' total factor productivity as well (Table A2, columns III-IV).

4.4.1 Dealing with potential endogeneity

Lastly, one could cast some doubts over the circular nature of the nexus between a firm productivity and the size of the agglomeration. This means that highly productive firms might "self-select", finding themselves to be more attracted to larger agglomerations, with this in turn affecting the extent of the agglomeration economies (Combes et al., 2012). As remarked in section 3.2, we believe that the multilevel approach, specifically including random effects at the level of the firm could control such self-selection mechanism. However, in order to reduce any potential concern, we try to run our basic model by considering the variable representing agglomeration as endogenous using a two-stages instrumental variables (2SLS/IV) approach. The selection of good instruments, correlated with the potential endogenous variable, but uncorrelated with our dependent variable, represents a big challenge in this context. In the end, we select two variables. The first is an index of city level productivity, which is constructed as a weighted average of all firms' productivity, excluding the firm i itself in order to avoid the risk of further endogeneity. The weights used are the shares of each firm on total employment. The second is a distance, calculated using city level coordinates, between the location of the firm and the capital city. In the context of African countries, in absence of reliable information on the size (in terms of total population or total area) of the urban areas, we can realistically assume that the capital cities are the larger in their countries and attract the bulk of the economic activities.

Results of the 2SLS/IV model, including the first stage regression, are reported in table A3 in the appendix. From the first stage regression, it can be shown that the two instruments are good predictors of the size of the agglomeration. Indeed, moving to the final output, the Sargan test of over-identification confirms that they are valid instruments. On the other hand, the Durbin-Wu-Hausman test rejects the null of endogeneity for our variable, reducing thus the concerns over the previous results. In any case, it can be observed that, also using this methodology, the results do not change consistently, once again confirming the robustness of our previous findings.

5. Conclusions

The idea that the development of African economies depends also on the performance of the domestic private sector is an old one. So far, however, only few evidence has been produced to show which factors do contribute to enhance the performance of domestic firms. Still, such existing evidence looks mostly at internal factors to the firm, and little attention has been given to the context in which they operate. This paper contributes to the existing literature by discussing the effects of sectorial and geographic agglomerations on firms' performance. This is done by exploring a large source of information at the firm level, the Africa Investor Survey, combined with an innovative methodology, the multilevel analysis, which allows to better understand how the external context in which a firm is embedded contributes to explain its performance.

One major contribution of our paper is to show that, taken together, the geographic environment and the industrial specialization contribute together to explain around 30% of the variance in domestic firms' productivity, with the location (country and city) accounting for more than half of this value. In addition, our paper provides a number of findings related to the agglomeration-performance nexus. We find that domestic SSA firms take advantage from the so-called urbanization economies, i.e. the co-location with other firms belonging to other industries. On the other hand, however, our analysis shows some forms of competitive effect arising from industry concentration, which is negatively correlated with firms' productivity. Both these effects are magnified when domestic firms are close to MNEs, and especially those from other developing countries.

Our results suggest some important implications for improving private sector performance in the context of SSA. The first is that the location of a firm has an influence on its performance and that the distribution of economic activities across well-defined geographic boundaries can contribute to the diffusion of positive externalities. Our results are consistent with existing evidence on the role of large cities in fostering economic development and productivity, given their potential to provide domestic firms with both economies of scale due to larger market size and other positive externalities due to the presence of urbanization economies (Jacobs, 1969). We find that this is especially true when the variety of economic activities is higher, as it can most likely to be associated with the exchange of information and pecuniary externalities. This is a not trivial argument in the case of many SSA countries, where the lack of economic diversification is often mentioned as one of the causes of backwardness.

As noted in previous research by Fafchamps and Soderbom (2014), proximity alone does not guarantee the transmission of spillovers. We find that this is especially true for local manufacturers, and show that such spillovers can be better absorbed in presence of linkages with other firms. On this respect, an interesting implication from our analysis is that, apparently, domestic companies receive a stronger contribution to enhance their productive efficiency from their local buyers rather than from foreign ones. This can be due to many reasons, including the well-known difficulties to adapt at the standards required by MNEs, which in most cases source only non-strategic resources from local firms, or the low absorptive capacities not allowing the full exploitation of the existing linkages (Morrisey, 2012).

Competition represents the other side of the story. Many scholars, supported by the experience of industrial districts in some developed countries, view competition as a key driver of firms' performance (Porter, 1990). The context of SSA is clearly different, and in fact we find contrasting findings. More specifically, we show evidence that an increase in the number of firms in the same industry (and/or producing the same product) is negatively correlated to productivity, and this is especially true when there are few bigger firms concentrating the market. Such result can be read together with the finding that the presence of MNEs from other Southern countries, sharing more similar production capacities, has a strong negative effect on domestic firms' productivity. Thus, one can assume that scale economies and the reduction of margins are significant strategies to face competition in such markets, potentially leading to a restructuring of smaller and less competitive firms, and also to their exit from the market. Even if we cannot control directly for firms' exit, this latter hypothesis is confirmed by some evidence on the so-called "crowding-out" of local firms following the entry of Chinese competitors (Kaplinsky and Morris, 2009).

References

Aitken, B.J. and Harrison, A.E. (1999) Do Domestic Firms Benefit from Direct Foreign Investment? Evidence from Venezuela, *American Economic Review*, 89(3): 605-618.

Alcacer, J., Chung, W., Hawk, A., and Pacheco-de-Almeida, G. (2013) Applying Random Coefficient Models to Strategy Research: Testing for Firm Heterogeneity, Predicting Firm-Specific Coefficients, and Estimating Strategy Trade-Offs, Harvard Business School Strategy Unit Working Paper N. 14-022

Amighini, A. and Sanfilippo, M. (2014) Impact of South-South FDI and Trade on the Export Upgrading of African Economies, *World Development*, 64: 1-17.

Andersson, M. and Loof, H. (2009) Agglomeration and Productivity – evidence from firm-level data, CESIS Working Paper Series N. 170.

Becattini G. (1990) The Marshallian industrial district as a socio-economic notion. In Pyke F., Becattini G., Sengenberger W. (Eds) *Industrial District and Inter-firm Cooperation in Italy*, pp. 37–51. International Institute for Labour Studies, Geneva.

Chhair, S. and Newman, C. (2014) Clustering, competition, and spillover effects – evidence from Cambodia, UNU-Wider Working Paper N. 2014/65.

Clarke, G.R.G (2012) Manufacturing Firms in Africa, in Dinh, H.T. and Clarke, G.R.G. (Eds.) *Performance of Manufacturing Firms in Africa An Empirical Analysis*, the World Bank: Washington, DC.

Combes, P.P., Duranton, G., Gobillon, L., Puga, D. and Roux, S. (2012) The Productivity Advantage of Large Cities: Distinguishing Agglomeration from Firm Selection, *Econometrica*, 80(6): 2543-2594.

Crespo, N. and Fontoura, M.P. (2007). Determinant Factors of FDI Spillovers – What Do We Really Know? *World Development*, 35(3): 410-25.

Ellison, G., Glaeaser, E. and Kerr, W.R. (2010) What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns, *American Economic Review*, 100: 1195-1213.

Fafchamps, M. and Hamine, S.E. (2004) Firm Productivity, Wages, and Agglomeration Externalities, CSAE Working Paper N. 2004-32

Fafchamps, M. and Soderbom, M. (2014) Network Proximity and Business Practices in African Manufacturing, *the World Bank Economic Review*, 28(1): 99-129.

Farole, T. (2013) *The Internal Geography of Trade - Lagging Regions and Global Markets*, the World Bank: Washington, DC.

Farole, T. and Winkler, D. (2013) *Making Foreign Direct Investment Work for Sub-Saharan Africa - Local Spillovers and Competitiveness in Global Value Chains*, the World Bank: Washington, DC.

Foster-McGregor, N., Isaksson, A. and Kaulich, F. (2013) Importing, exporting and performance in sub-Saharan African manufacturing firms, *Review of World Economics*, DOI 10.1007/s10290-013-0177-y

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

Gill, I. and Goh, C.C. (2010) Scale Economies and Cities, *The World Bank Research Observer*, 25(2): 235-262.

Gorg, H. and Greenaway, D. (2004). Much Ado about Nothing? Do Domestic Firms Really Benefit from Foreign Direct Investment? *World Bank Research Observer*, 19(2), 171-97.

Gorg, H. and Seric, A. (2013) With a little help from my friends: Supplying to multinationals, buying from multinationals, and domestic firm performance, Kiel Working Papers N. 1867.

Grossman G.M. and Helpman E. (1991) *Innovation and Growth in the Global Economy*. MIT Press, Cambridge MA.

Henderson, J.V. (2003) Marshall's scale economies, *Journal of Urban Economics*, 53: 1-28.

Howard, E., Newman, C. and Tarp, F. (2012) Measuring Industry Agglomeration and Identifying the Driving Forces, Learning to Compete Working Paper N. 3, Brookings Institute.

Howard, E., Newman, C., Rand, J. and Tarp, F. (2014) Productivity-enhancing manufacturing clusters – evidence from Vietnam, UNU-Wider Working Paper N. 2014/71.

Iacovone, L., Ramachandran, V. and Schmidt, M. (2014) Stunted Growth: Why Don't African Firms Create More Jobs? Center for Global Development Working Paper 353.

Krugman, P. (1991) Increasing Returns to Scale and Economic Geography, *Journal of Political Economy*, 99(3):483–99.

Lipsey, R.E. and Sjoholm, F. (2011). South-South FDI and Development in East Asia. *Asian Development Review*, 28(2), 11-31.

Jacobs, J. (1969). *The Economy of Cities*, Random House.

Marshall, A. (1920). *Principles of Economics*. London: MacMillan.

Maas, C.J.M., and Hox, J.J. (2004) The influence of violations of assumptions on multilevel parameter estimates and their standard errors, *Computational Statistics and Data Analysis*, 46: 427-440.

McCormick, D. (1999). "African Enterprise Clusters and Industrialization: Theory and Reality." *World Development*, 27 (9): 1531–51.

Melitz, M. (2003). The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica*, 71(6), 1695-1725.

Morrisey, O. (2012). FDI in Sub-Saharan Africa: Few Linkages, Fewer Spillovers. *The European Journal of Development Research*, 24(1), 26–31.

Narula, R. and Driffield, N. (2012). Does FDI Cause Development? The Ambiguity of the Evidence and Why it Matters. *European Journal of Development Research*, 24(1), 1-7.

Perez-Villar, L. and Seric, A. (2014) Multinationals in Sub-Saharan Africa: Domestic linkages and institutional distance, Kiel Working Paper, No. 1893.

Porter, M.E. (1990) *The Competitive Advantage of Nations*. New York: The Free Press.

Rabe-Hesketh, S., and Skrondal, A. (2008) *Multilevel and Longitudinal Modelling Using STATA*, STATA Press.

Romer., P. (1990) Endogenous technological change, *Journal of Political Economy*, 98: 71-102.

Siba, E., Soderbom, M., Bigsten, A. and Gebreeyesus, M. (2012) Enterprise agglomeration, output price, and physical productivity: firm-level evidence from Ethiopia, UNU-Wider Working Paper N. 2012/85.

Snijders, T.A.B., and Bosker, R.J. (1999) *Multilevel analysis*, London: Sage Publications.

UNIDO (2012) African Investor Survey, United Nations Industrial Development Organization: Vienna.

Van Biesebroeck, J. (2005) Firm Size Matters: Growth and Productivity Growth in African Manufacturing, *Economic Development and Cultural Change*, 545-583.

Van Oort, F.G., Burger, M.J., Knoben, J. and Raspe, O. (2012) Multilevel Approaches and the Firm-Agglomeration Ambiguity in Economic Growth Studies, Tinbergen Institute Discussion Paper N. 2012-014/3

World Bank (2009) *World Development Report*, the World Bank: Washigton, DC.

Yehoue, E.B. (2005) Clusters as a Driving Engine for FDI, IMF Working Paper N. 05/193.

Yoshimo, Y. (2012) *Industrial Clusters and Micro and Small Enterprises in Africa*, the World Bank: Washigton, DC.

Appendix

Figure A1. Top agglomerations (total N. of firms)

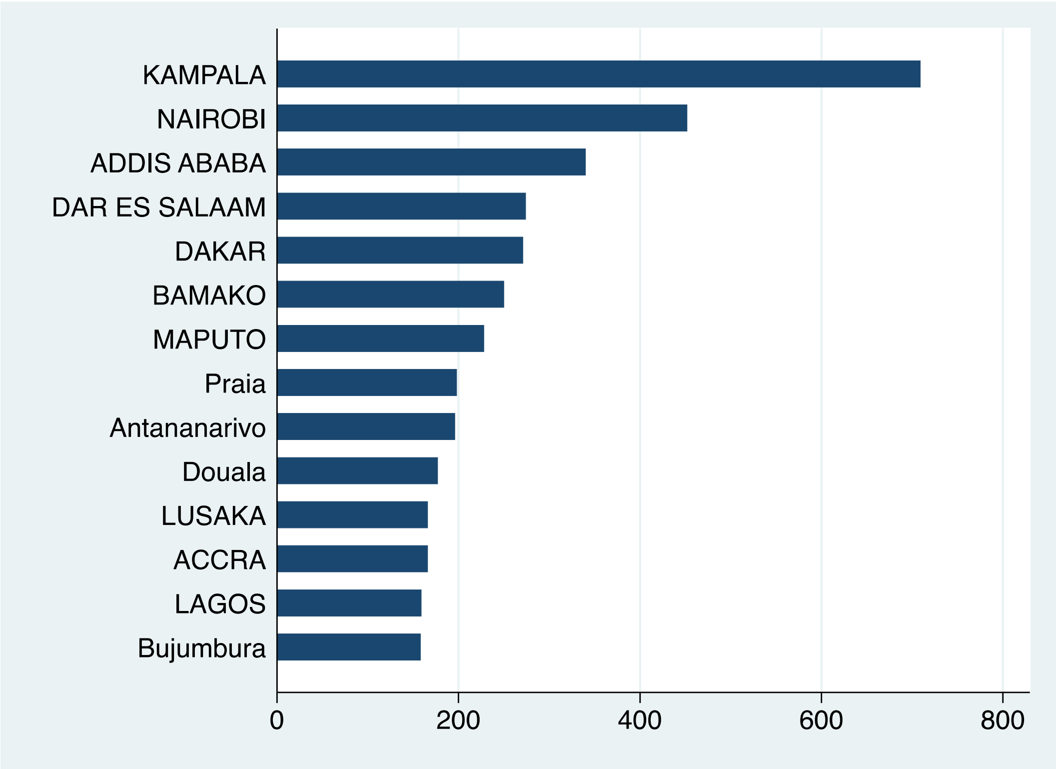


Table A1. Descriptive statistics

Variable	Description	Obs	Mean	Std. Dev.	Min	Max
LP	Labor productivity, log	3612	2.07	1.60	-5.83	10.20
Size	Size classes (Small=1, Medium=2, Large=3)	3825	1.73	0.86	1	3
Age	Age, log	3826	2.64	0.78	0	5.09
Exp	Dummy, 1 if exporting	3426	0.18	0.38	0	1
Family	Dummy, 1 if family owned	3865	0.40	0.49	0	1
Skill	Share of skilled workers on total	3723	23.00	21.63	0	100
R&D	Dummy, 1 if R&D expenditures>0	3865	0.64	0.48	0	1
N_c	Firms in the same city	3865	215.19	200.00	0	709
$N_{j,c}$	Firms in the same city and industry	3865	11.48	13.40	0	60
$N_{x,c}$	Firms in the same city and different industry	3865	203.71	191.34	0	709

Table A2. Results, robustness checks

	(I) lab_prod	(II) lab_prod	(III) tfp	(IV) tfp
Size	0.171***	0.174***	0.135***	0.135***
	(0.046)	(0.045)	(0.045)	(0.045)
Age	0.141***	0.142***	0.090**	0.089**
	(0.034)	(0.034)	(0.035)	(0.035)
Exp	0.481***	0.484***	0.336***	0.336***
	(0.072)	(0.072)	(0.066)	(0.065)
Family	-0.230***	-0.232***	-0.115**	-0.115**
	(0.074)	(0.074)	(0.051)	(0.051)
skill	0.007***	0.007***	0.005***	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)
R&D	0.029	0.029	0.021	0.022
	(0.065)	(0.065)	(0.064)	(0.063)
$L_{x,c}$		0.005**		
		(0.002)		
$L_{j,c}$		-0.017		
		(0.012)		
L_{c}	0.004**			
	(0.001)			
N_{c}			0.000*	
			(0.000)	
$N_{j,c}$				-0.003
				(0.004)
$N_{x,c}$				0.001**
				(0.000)
Constant	1.061***	1.060***	6.011***	6.007***
	(0.234)	(0.231)	(0.273)	(0.268)
Obs.	3,281	3,281	3,257	3,257
Industry Effects	Yes	Yes	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: The variables "L" represents the total number of employees, and are expressed in thousands.

Table A3. Results, IV regression

	(I) First stage	(II) 2SLS
N_c		0.001**
		(0.000)
Size	1.424	0.174***
	(1.993)	(0.032)
Age	0.42	0.138***
	(2.091)	(0.034)
Exp	16.096***	0.500***
	(4.447)	(0.072)
Family	4.767	-0.235***
	(3.203)	(0.052)
skill	-0.014	0.007***
	(0.0773)	(0.001)
R&D	-7.613*	0.040
	(4.042)	(0.066)
dist_cap	-33.241***	
	(0.727)	
city_productivity	24.924***	
	(1.982)	
Constant	26.214	1.872***
	(19.414)	(0.293)
Observations	3,219	3,219
R-squared	0.8461	0.272
Sargan Chi2 (p-value)		0.1393
Durbin-Wu-Hausman (p-value)		0.1098

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Author contacts:

Marco Sanfilippo

European University Institute

Robert Schuman Centre for Advanced Studies

Via delle Fontanelle, 18

50014 San Domenico di Fiesole

Italy

Email: marco.sanfilippo@eui.eu

Adnan Seric

United Nations Industrial Development Organization

Wagramerstr. 5

1400 Vienna - Austria

Email: a.seric@unido.org