ECONOMIC RESEARCH FORUM

منتدى البحوث الاقتصادية

Working Paper Series

25TH ERF ANNIVERSARY

# FIRM PERFORMANCE AND AGGLOMERATION EFFECTS: EVIDENCE FROM TUNISIAN FIRM-LEVEL DATA

Mohamed Amara

Working Paper No. 1297

# FIRM PERFORMANCE AND AGGLOMERATION EFFECTS: EVIDENCE FROM TUNISIAN FIRM-LEVEL DATA[1]

Mohamed Amara[2]

**Working Paper No. 1297**

**February 2019**

We are grateful to an anonymous referee for helpful comments and suggestions.

Send correspondence to:
Mohamed Amara
University of Tunis
Mohamed.amara.isg@gmail.com.

---

[1] We gratefully acknowledge financial support from the Economic Research Forum (ERF) under the project "Structural Change, Resource Misallocation and Growth Dynamics in the MENA Region". The views expressed in this paper are those of the author and do not necessarily represent those of the ERF.
[2] Higher school of Economic and Commercial Sciences of Tunis, University of Tunis & MENA Youth Policy Center (MYPRC).

First published in 2018 by
The Economic Research Forum (ERF)
21 Al-Sad Al-Aaly Street
Dokki, Giza
Egypt
www.erf.org.eg

## Abstract

Using Tunisian manufacturing data between 1998 and 2004 and by referring to multilevel approach, this paper investigates the impact of agglomeration and individual characteristics on firm's performance. The empirical results show the importance of considering both regional and firm characteristics when examining firm performance. They also support the validity of self selection and learning-by-exporting hypotheses. Urbanization and localization effects are significant and positive for firm's export behavior, but only localization economies have a positive effect of firm productivity. However, the results of the quantile approach show that selection, rather than agglomeration economies in larger cities, better explain spatial productivity differences.


**Keywords:** Firm performance, Agglomeration effects, Firm selection, Multilevel, Tunisia.

**JEL Classifications:** D24, L25, R12.

## 1. Introduction

The economic geographical literature has forged strong linkages between geographical space and the performance of regions and firms, largely in terms of attributes such as export performance, competitiveness, and innovation. The qualities of specific places, such as proximity to export markets, good infrastructure, benefits of agglomeration, and the availability of skilled workers are largely considered as potential determinants of firm performance in contemporary research (Rodríguez-Pose and Hardy 2017). The 2009 World Development Report *Reshaping Economic Geography* and the 2011 Middle East and North Africa (MENA) flagship report *Poor Places, Thinking People* emphasize the importance of government and institutional factors that are instrumental in shaping the evolution of disparities between regions. The main message of those reports is that economic growth is a function of density, distance and agglomeration, and thus will likely be unbalanced depending on those variables (World Bank 2014).

The economic activities as well as the per capita income levels differ significantly across geographical areas in MENA countries. Tunisia is a good example of this, although there are, of course, significant differences among MENA countries. In fact, although the Tunisian economy has shown relatively robust aggregate growth of about 5 percent per year over the past decade before the 2010 beginning of the revolution, wide-spread inequalities between coastal and inner regions persist and grow worse. 85% of the country's GDP is generated by the coastal area and 63% in the three agglomerations of Tunis, Sfax and Sousse (World Bank 2014). The share of manufacturing firms reaches 84% in the coastal area, which absorbs 88% of the manufacturing jobs in 2010 against only 12% for the non-coastal areas (Amara and Ayadi 2014). Moreover, 85% of the economically active population estimated at 3.9 million in 2014, are employed in the service and manufacturing sectors essentially localized in urban areas. The effects of those external factors (proximity of population concentrations, labor pool and infrastructure availability, etc) involve shifts of a firm's production or cost curves. That is, factor external to a particular firm but associated with firm clustering increase firm productivity, implying more output for a given amount of inputs or less input cost to produce a given amount of output (Cohen and Paul 2009). In addition, the effects of such agglomeration involve various aspects of economic performance from clustering such as enhanced innovation (Audretsch and Feldman 1996), higher input (labor or capital) demand or price (Moretti 2004), greater productivity (Henderson 2003) and reduced costs (Cohen and Paul 2005).

Bearing in mind these considerations, this paper tries to estimate the impact of agglomeration on firm-level performance (in terms of total factor productivity (TFP) and export performance) in manufacturing sector based on micro and regional-level data. The contribution of this work lies in the following two aspects: First, this is the first study at the micro-level, which decomposes the effect of firm characteristic versus agglomeration in a rigorous framework. Since it relies on the use of firm-level data and is among the few studies that explore agglomeration effects in Tunisian economy, it offers valuable contributions to the industrial economics literature in MENA region and in developing countries in general. Second, to the best of our knowledge, this is the first study in MENA region, which distinguishes agglomeration (larger cities promote interactions that increase firm performance) from firm selection (larger cities allow only the most productive firms to survive). This represents an innovation in the empirical literature.

Through this study, we seek to answer the following questions: What is the effect of agglomeration economies on firm performance (total factor productivity (TFP) and exporting behavior)? How important are such agglomeration effects? And are they likely to be relevant to economic development policy in Tunisia? If so, what sorts of policy instruments are relevant in exploiting these advantages for more inclusive and sustainable growth? Using an unbalanced panel of 8200 firms in 11 manufacturing industries over the period 1998-2004, we assess how agglomeration determines firm performance. We apply a multilevel analysis to explain the differences in the TFP and export participation by providing a clear distinction between firm (micro level) and agglomeration effects (macro level). In this respect, multilevel analysis yields a decomposition of TFP variance and provides highly informative outcome related to the quantitative measure on 'how much' regional and individual factors explain of TFP heterogeneity (Aiello and Ricotta 2016). In addition, it is possible to show how agglomeration effects translate into individual behavior (productivity and export). Multilevel models provide also a useful empirical framework to the micro-macro problem known as "ecological fallacy" or "cross-level fallacy" (Robinson 1950, Alker 1969). They allow us to incorporate unobserved heterogeneity into the model via random intercepts and allowing relationships to vary across contexts via the inclusion of random coefficients (Snijders and Bosker 1999, Van Oort et al. 2012). We also use the same estimation approach set forth Combes et al (2012) to separate out selection from agglomeration effects.

The rest of the article is organized as follows. The next section reinforces the motivation of the paper in a brief theoretical foundation on the relationship between agglomeration and firm performance. Section 3 discusses the data used and the model specification details, as well as the basic definitions of the variables and the expected signs. Section 4 presents and discusses the empirical results and section 5 performs robustness analysis. Section 6 concludes with some policy implications.

## 2. Economic agglomeration and industrial policies in Tunisia

### 2.1. Agglomeration and firm performance: theoretical foundations

In the literature dealing with agglomeration economies and competitiveness of firms, scholars argue that knowledge spillovers, informal relations and social interactions have an effect, not only on firm performance but also on economic growth (Audretsch and Feldman 2004, Henderson 2007, Henderson and Wang 2007). In the literature reviews, productivity is considered at two different levels: the micro level and the macro level. In a micro-level, giving the increasing availability of individual firm data, a growing number of studies have tried to identify what factors influence the productivity of firms. Most of these studies have been carried in advanced economies such as the United States, Germany, France, Italy and the United Kingdom (Keller and Yeaple 2009, Wagner 2007, Martin et al 2011, Parisi et al 2006, Wakelin 2001). Firms are naturally influenced by their own attributes and resources (also known as internal factors) such as competencies, knowledge and human capital (Backman

2014). Human capital can impact firm's performance through several mechanisms (Ballot et al 2001): (1) a firm who has substantial human capital will make better decisions than its rivals with lower human capital; (2) innovation will be stimulated by the quality and training of the personnel in the R&D department; (3) learning-by-doing is also higher if workers have high human capital. The empirical literature also suggested that firm size and firm age have a positive impact on productivity. Indeed, firm size largely determines a firm's resource base, competencies and scale advantages. Due to internal economies of scale that reduce the per-unit costs over the number of units produced, efficiency advantages emerge from larger firm sizes, while small firms have to overcome these disadvantages (Raspe and Van Oort 2011a,b). This finding was confirmed recently by Ayyagari et al (2011) using the World Bank Enterprise Surveys for developing countries to show that large firms are typically more productive than small firms. Still remaining at micro-level, the concept of knowledge spillovers and firm productivity has received increasing interest over the past decades because the external environment is a critical factor affecting the firms' performance (Henderson et al 1995, Audretsch and Feldman 1996, Ellison and Glaeser 1997, Rice et al 2006). Recently theoretical developments have attempted to open the 'black box' of knowledge spillovers and to explain how these spillovers work at the micro level (Duranton and Puga 2004, Henderson 2007). In other words, they seek to understand how local interactions, peer effects, spatial relationships and social networks lead to better firm performance, such as productivity levels (Ciccone and Hall 1996, Cingano and Schivardi 2004).

Based on the work of Marshall (1920), Duranton and Puga (2004) list the following three mechanisms or sources of micro-foundations of agglomeration economies: input sharing, matching or labor market pooling and learning or knowledge spillovers. The first refers to benefits that arise due to larger variety of input suppliers and a deeper division of labor force at the same region. The second refers to matching models by which agglomeration improves the expected quality of matches and increases the probability of matching. It involves a concentration of workers that reduces the risks and costs of searching for workers or jobs. The third mechanism refers to the fact that agglomeration offers opportunities for the generation, the diffusion and the accumulation of knowledge spillovers among firms and workers. These mechanisms provide arguments for why productivity may be enhanced in urban agglomeration economies, and thus reasons why they persist.

Taking the analysis to the regional-level, the majority of current research estimates production function at the aggregate level (such as state, city or country) (Rice et al 2006, Ke 2010). In this context, productivity differences and regional convergence are issues of intense theoretical and empirical research since the development of New Growth theory and New Economic Geography (NEG). These studies suggest that regional gap in productivity can be attributed to regional differences in various factors such as education endowment, foreign direct investment (FDI), producer's and customer's market accessibilities, and agglomeration economies. These factors contribute to the agglomeration of firms in urban areas. Krugman (1991) showed, in addition, that decline in transport costs, increases of economies of scale, and mobility of the specialized labor reinforce agglomeration of firms and increase regional disparities. Concerning the agglomeration economies' effects and moving from the pioneering contributions of Glaeser et al (1992) and Henderson et al (1995), three types of agglomeration externalities have been widely investigated: specialization (or Marshall-Arrow-Romer externalities[3], MAR for short), diversification (or Jacobs externalities) and urbanization externalities. MAR externalities stem from intra industry knowledge spillovers. They support the hypothesis that local industrial specialization is an effective way to increase knowledge spillovers between firms and therefore boost the growth process and productivity. In addition, sharing the same field of activity facilitates interactions among firms having the same problems and the same concerns. In Marshallian tradition, MAR externalities can be attributed to three important sources (Neffke et al 2011): labor market pooling, input-output linkages, and intra-industry knowledge spillovers. The third source seems to be winning great importance in the explanation of the spatial polarization of economic activity. In contrast to MAR externalities, Jacobs' externalities arise from local diversity, external to industry or sector. Local diversification within an urban region fosters innovation and result in cross fertilization of ideas that born in the exchange process that occurs between different fields of knowledge (Jacobs 1969). Urbanization externalities are associated with scale economies' benefits arising from large and diversified urban areas (Cainelli et al 2016), where agents can share research centers, universities, high-tech services and transportation systems.

Since the work of Glaeser et al. (1992) and Henderson et al. (1995), empirical studies on the relationship between externalities and firm's performance have been an important development. Glaeser et al. (1992) examined the impact of the externalities on the local growth in 170 of the largest U.S cities between 1956 and 1987. The authors find that only a diversified industrial city has a significant impact, revealing the presence of Jacobs' externalities rather than MAR externalities. In contrast, Henderson et al. (1995) find evidence of MAR externalities for all industries whereas Jacobs's externalities are limited to high-tech industries. Using firm-level at Italian local labor systems (LLS), Cingano and Schivardi (2004) find that industrial specialization and scale indicators affect total factor productivity growth positively. Based on 67 reviewed articles on localization versus urbanization debate, Beaudry and Schiffauerova (2009) argue that the level of industrial or geographical aggregation and the choice of performance measures (specialization and diversity indicators) are the main causes for the lack of resolution in the debate.

Some other works have used a multilevel modeling which allows micro and macro levels to be modeled simultaneously (Aiello et al. 2014, Backman 2014, Raspe and Oort 2011a,b, Fazio and Piacentino 2010, Amara and Thabet 2016). Multilevel modeling can also reduce the ambiguity surrounding the agglomeration-firm performance relationship and consider spatial, sectorial and cross-level heterogeneity (Fazio and Piacentino 2010). Aiello et al. (2014) find that both micro and macro factors affect total factor productivity of Italian manufacturing firms. More specifically, they show that regional endowment of infrastructure, the efficiency of local administrative and the investments in R&D exert a positive effect on firms' performance.

[3] In reference to contributions of Marshall (1920), Arrow (1962), Romer (1990).

## 2.2. Industrial policies in Tunisia

Multiple reasons may explain the persistence of the regional inequality in Tunisia but two traits are pervasive. First, Tunisia inherited a considerable infrastructure for production and distribution facilities concentrated in coastal regions, which had been set up by the French protectorate. Second, the inadequacy of some industrial policies would have contributed to exacerbate the problem. During the 1970s and after the failure of the cooperative experiment (1964-1969), the country adopted an industrial policy focused on export promotion and private investment. In order to support this policy and provide technical assistance to investors, the government created in 1973 the *Agence pour la Promotion de l'Indistrie* (API) and the *Fond pour la Promotion et Décentralisation Industrielle* (FOPRODI) (Cammett 2007).[4] From 1970 to 1979, the Tunisian economy performed strongly. The GDP growth averaged 7.4 percent a year over the decade, aided by the oil exports and the high levels of investment (specifically the foreign direct investment through fiscal incentives). More than 500 foreign firms established their production units under the 1972 law between 1973 and 1978 (Jelili and Goaied 2010). Most of these firms are concentrated in industrial zones along the coast, from the city of Bizerte, on Tunisia's northern coast, to Sfax on the eastern coast. In this context, the FOPRODI was created in 1974 to promote decentralization, but no concrete results were obtained.

After the period of deteriorating economic performance between 1980 and 1984, Tunisia has adopted in 1986 a Structural Adjustment Program (SAP) to speed up privatization and deepen integration with European markets (Diop 2008). To achieve such goals, Tunisia committed more consciously to industrial clustering as a means to boost productivity levels of both firms and the labor force. These reforms stabilized the economy and improved the external position, boosting exports by an average 6.6 per cent a year at constant prices between 1987 and 2001, while annual inflation fell from 8.1 per cent to 1.9 per cent over the same period[5]. The privatization program led to the full or partial privatization of nearly 160 state-owned companies (Ayadi and Mattoussi 2014).[6] Unfortunately, industrial clustering has affected the spatial structure of economic activities by increasing inequalities in economic performance and employment opportunities between coastal and non-coastal areas and may therefore have played a role in the deepening regional disparities.

In the 1990s, the Tunisian state also created sector-specific technical support centers, such as the *Centre Technique du Textile* (CETTEX), which provides advice and expertise for the textile industry and supports foreign partners in their search for business opportunities.[7] The government also created the *Mise à Niveau* program (PMN) in 1995 to enhance competitiveness and accelerate the business processes of modernization. Moreover, the governorate installed two free trade zones in Bizerte (60 km north of Tunis) and in Zarzis (450 km south of Tunis) to offer more favorable environment for foreign investors.

[4] Cammett, M. (2007). Business–government relations and industrial change: The politics of upgrading in Morocco and Tunisia. *World development*, *35*(11), 1889-1903.
[5] African Development Bank, & Development. Development Centre. (2003). *African economic outlook*. OECD.
[6] Ayadi, M., & Mattoussi, W. (2014). *Scoping of the Tunisian economy*, WIDER Working Paper, No. 2014/074.
[7] Fair wear foundation. Tunisia country study 2015.

## 3. Data and Research Methodology

### 3.1. Data

The main data set used for this study is the National Annual Survey Report on Firms (NASRF) carried out by Tunisian National Institute of Statistics (TNIS) over the period 1998-2004. The TNIS collects annual unbalanced-sheet data on a sample of 5000 firms from the Tunisian Business Register (RNE) covering almost all formal sectors (firm that has employed six or more people), out of which about 2000 responded to the questionnaire. The information is organized in four areas: information relating to the identification of the firm and its general characteristics; information relating to production factors (employees, investments, depreciation and amortization); information relating to production (purchases, inventories, production and sales) and information relating to accounting. The data has been cleaned by removing outliers and missing values. As only manufacturing firms were used, the service sectors, the extractive sectors and so forth were not considered. Thereby, our empirical analysis is based on an unbalanced panel consisting of a sample of about 8200 firms from 11 manufacturing sectors over 7 years (about 1170 firms in average observed by year).

Additionally, we use regional and sector-level variables from the Tunisian Business Register (RNE) for the period 1998-2004 collected and continuously updated by the TNIS. The Tunisian Business Register is an annual census combining information from many different sources n such as the National Social Security Fund (CNSS), the Tunisian Customs, The Ministry of Finance and the Agency for the Promotion of Industry and Innovation (APII). Data from the Ministry of Finance contain information about opening statements and therefore constitute the most comprehensive source for identifying companies in the country. Data from the CNSS provide information about establishments such as workforce and its characteristics, training level, working conditions and the level of income from work for both employees and employers. According to the RNE, there were 654,524 private firms in 2013, a net increase of about 299,612 since 1996. More than 88% of all private firms were one-person businesses, and almost 97% were microenterprises with fewer than six workers. Only 2.34% were small enterprises (6-49 employees), 0.39% were medium enterprises (50-199 employees), and only 0.12% (779 firms) had 200 or more employees.

### 3.2. Measures at the firm level

All firm-level variables used in this study and presented below are drawn from the National Annual Survey Report on Firms.

#### 3.2.1. Performance

We use firm-level productivity and firm export participation (or firm's export probability) as our performance measures. The Olley and Pakes (1996) approach was employed to measure the firm productivity (more details will be presented in the methodology section). We use two variables to measure the extensive margin of exports (the export participation). The first one is a dummy variable which distinguishes between exporting and non-exporting firms, and the second one distinguishes between fully exporting firms from others (partially exporting and non-exporting firms). In order to distinguish extensive margin from intensive margin of exports, we added the share of exports over the total sales as a new dependent variable.

There are two hypotheses in the literature regarding the relationship between exporting and productivity. The first is related to self-selection and argues that only high-productivity firms will be able to enter the export market. Firms that export incur extra costs (transportation costs, costs of managing foreign networks, etc) and without prior high productivity, they cannot afford to export their products and services (Roberts and Tybout 1997, Bernard and Jensen 1999, Harris and Li 2008). The second hypothesis is the learning-by-exporting mechanism assuming that firms improve their productivity by participating in global markets (Aw et al. 2000). It was argued that exporting firms benefit from foreign competition and technology which make it possible for them to improve their productivity as well as to exploit economies of scale (Clerides et al. 1998, Aw et al. 2000). Most empirical studies have found strong evidence to support the self-selection hypothesis and weak evidence on learning-by-exporting (Kox and Rojas-Romagosa 2010).

#### 3.2.2. Firm age and firm size

It is largely believed that firm size and firm age positively impact firm performance (Jovanovic 1982). Firm age might capture the extent of a firm's learning experience over time (Raspe and Van Oort 2011). Older firms learn from their own experience and previous operations and are generally considered to be more efficient than younger firms. Furthermore, experience and learning over time can also have a significant impact on the decision and export intensity of firms. On the other hand, it can be hypothesized that firm performance decline with age that can be linked to inflexibility and a failure to change strategy and behavior (Raspe and Van Oort 2011).

With respect to the impact of size, the empirical literature agrees in considering size to be a main source of heterogeneity in firm's productivity. Large firms are generally considered to be more productive than smaller-sized firms. Thus, despite the fact that smaller firms have the advantage of flexible management and lower response time to market changes, larger ones find it easier to benefit of economies of scale and to access to credits (De and Nagaraj 2014). Additionally, and given that exporting is often associated with high-level productivity, large firms are more likely to export their products compared to smaller firms. This finding has significant policy implications and it is particularly important for developing countries that have a large proportion of small and medium firms.

### 3.3. Measures at the regional level

All regional variables used in this study and presented below are drawn from Tunisian Business Register.

#### 3.3.1. Localization economies

Marshall (1890) was the first to make the distinction between economies of scale internal to the firm and external economies of scale arising from the co-presence of many firms of the same sector in a region. Marshall supported the hypothesis that local industrial specialization is an effective way to increase firm performance. He identified three distinct sources through which external economies arise: knowledge spillovers, labor market pooling, and inter-firm linkage. The presence of specialized firms in a region may lead to the sharing of supplies and the matching between employers and employees. In addition, the concentration of an industry in a region promotes the generation and diffusion of ideas and technology between firms and facilitates innovation in that particular industry within that region (Duranton and Puga 2004, Beaudry and Schiffauerova 2009). The importance of technological spillovers was later also underlined by Arrow (1962) and Romer (1990), so that localization economies are also known as Marshall-Arrow-Romer externalities or MAR externalities in dynamic context (Glaeser et al. 1992). Porter (1990), like MAR, argues that specialization is beneficial for both firm in the same sector and the city in which it was located. He contends, however, that local competition speeds up the adoption of technology and drives firms to innovate. Thus, Porter considers that specialization and local competition are two key factors for success and growth as well.

In empirical literature, several indicators have been used to measure the MAR externalities, which can be classified into two groups (Fracasso et al. 2018): size-based indexes (e.g., industry employment, number of industry plants, employment related industries) and share-based indexes (e.g., share of industry in the region, location quotient (LQ)). According to the meta-analysis provided by Beaudry and Schiffauerova (2009), almost half of the reviewed studies employed the share-based indexes. In addition, Beaudry and Schiffauerova (2009) found that the location quotient and own-industry employment (belonging to group of share-based indexes), which together account for 75% of independent variables used in 67 reviewed articles, are the most common MAR externalities indicators. The popularity of the LQ is likely due to its appearance in the work of Glaeser el al. (1992), where the authors expressed the idea that the degree of specialization (that captures intensity and density of interaction among firms) may better represent the potential for MAR externalities than current size of an industry (see Fracasso et al. 2018 for theoretical and empirical motivations that may justify the adoption of the LQ to capture MAR externalities). In line with Glaeser et al. (1992) we use the LQ index, called also Hoover-Balassa coefficient or specialization index, to measure specialization. It is calculated as the fraction of industry employment in a region relative to the national share.

$$\text{specialization} = \frac{\text{employment of sector s in region/Total employment in region j}}{\text{Total employment in sector s /total employment}} = \frac{\text{emp}_{\text{sjt}}/\sum_{\text{s}} \text{emp}_{\text{sjt}}}{\sum_{\text{j}} \text{emp}_{\text{sjt}}/\sum_{\text{s}}\sum_{\text{j}} \text{emp}_{\text{sjt}}} \qquad (1)$$

where $\text{emp}_{\text{sjt}}$ is the employment of sector s in region j at time t. Values above 1 imply that a certain industry is overrepresented at a particular region, as compared to the average situation at the national level.

### 3.3.2. Urbanization economies

In contrast to MAR externalities, Jacobs (1969) argues that knowledge can spillover between complementary rather than similar industries. In other words, regions with a diversified economic structure will tend to grow faster than specialized ones. So, agglomeration economies may not necessarily relate to specialization, but can therefore arise from the overall size and diversity of economic activity in a given region (Jones 2017). In this case, they are defined as Jacobs externalities and also known as urbanization (diversification) economies (Cohen and Paul 2009). According to Jacobs externalities, the interaction between clustered firms belonging to different industries may foster innovation and local economic growth through the cross-fertilization of ideas that born in the exchange process that occurs between different fields of knowledge (among others, see Beaudry and Schiffauerova 2009 and Melo et al. 2009). In addition, firms in such industrial atmosphere can benefit from economies of scale arising from large and diversified agglomeration, where it is easy to share research centers, universities, high-tech services and transportation systems (Cainelli et al. 2016).

Indicators of Jacobs externalities can be classified into two categories (Beaudry and Schiffauerova 2009). The first category is based on industrial diversity, while the second one is based on market size. The Hirschman-Herfindahl index (HH), approximated as the squared shares of employment in a given region and sector with respect to all other industry employment (Henderson, 1997), is the most commonly used indicator in the empirical literature. It is also used in modified forms, as inversed HH index or 1 minus the HH index (Beaudry and Schiffauerova 2009). Some other studies use the total employment in the region or total population in the region to approximate Jacobs externalities. These indicators, however, capture market size rather than diversity because they derive from the specific industrial composition of the region (Beaudry and Schiffauerova 2009). Despite the extensive literature on this subject, the choice of diversity measurement remains unclear and depends on the independent variable to be used and the methodology to be employed. In our study, we use the inverse of the HH, which is the most commonly used index, to capture the diversity (Duranton and Puga 2000). For each region j this sums over all sectors the square of each sector's share in local employment. Formally, the inverse of the HH index or diversity index is given by

$$\text{Diversity index}_{jt} = \frac{1}{HH_{st}} = \frac{1}{\sum_s (S_{jst})^2} \qquad (2)$$

Where $S_{jst}$ indicates the local employment's share of sector s for region j at time t. The diversity index increases as activities in the region under consideration become more diverse and it is equals to 1 if the economic activity is fully concentrated in one region. A relative diversity index has been proposed by Duranton and Puga (2000) in order to consider the difference between each sector's share in local employment and its share in national employment. Formally, the relative diversity index is given by

$$\text{Relative diversity index}_{jt} = \frac{1}{\sum_s |S_{jst} - S_{st}|} \quad (3)$$

where $S_{jst}$ is the share of industry s in region j at time t and $S_{st}$ represents the share of industry s in the national level during the same time period.

Although an extensive literature has been devoted to analyze the relationship between firm performance and MAR and Jacobs externalities, the debate about which one is better remains unresolved. Glaeser et al. (1992) show for example that only a diversified industrial city has a significant impact on local growth in 170 of the largest US cities between 1956 and 1987, revealing the presence of Jacobs' externalities rather than MAR externalities. In contrast, Henderson et al. (1995) found evidence of an impact of MAR externalities for all industries whereas Jacobs externalities are limited to high-tech industries. Combes (2000) provided different results for the case of France, where industry employment is negatively influenced by both specialization and diversity, while service employment only benefits from diversity. Combes et al. (2004) explain the divergence in the empirical results by the difference in their methodologies (plant-level versus regional studies; panel data versus cross section analyses and employment versus productivity regressions). Moreover, studies can widely differ among countries (developing versus developed countries) and periods or between geographic areas in the same country (Neffke et al. 2011).

Table A.1 in the appendix describes further characteristics of the firm-level and regional-level data. Firm's log productivity is on average equal to 6.97 with a standard deviation of 0.75. Chemical firms are more productive than those in other manufactures. Table A.1 shows in addition that firms with equal or more than 200 employees are more productive and have more chance to export than smaller firms with 50 or less employees.

### 3.4. Methodology

To achieve our objective of studying the effects of agglomeration economies on firm productivity and underlying causal mechanism linking productivity to exporting, we propose two hierarchical models: productivity model and export model. For the productivity model, we extend the works of Bernard and Jensen (1999, 2004) and Aiello et al (2014) by using simultaneously micro data at the level of firms and macro data at the level of governorates to investigate the relationship between export activities, productivity and agglomeration economies empirically. Suppose a firm $i$ produces output $Y_i$ using the following Cobb-Douglas production function:

$$Y_i = T_i K_i^{\beta_K} L_i^{\beta_L}, \; i = 1, \dots, N \qquad (4)$$

Where $K_i$ and $L_i$ denote the capital stock and the labor-force (in terms of employees), respectively. $\beta_K$ and $\beta_L$ are the parameters of interest, $T_i$ is the total factor productivity (TFP) and N the total number of firms. Over a panel data and proceeding by a logarithmic transformation of equation (4), the estimating equation is giving by:

$$y_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + u_{it} \qquad (5)$$

$$u_{it} = \omega_{it} + \eta_{it}$$

Where $y_{it}$ is the log of output (value-added) from firm i at time t, $k_{it}$ the log of its capital and $l_{it}$ the log of its labor input; the $\beta_k$ and $\beta_l$ are coefficients to be estimated (interpreted also as output elasticity relative respectively to capital and labor). The error term $u_{it}$ consists of two components: the stochastic term $\eta_{it}$ and the productivity $\omega_{it}$. The $\eta_{it}$ represent shocks to production that are uncorrelated with the input choices and not observable by firms before making their input decisions at time t. In contrast, the $\omega_{it}$ represent the productivity shocks that are observed by firms and act as variables (such as the managerial ability of a firm, expected down-time due to machine breakdown, expected defect rates in a manufacturing process) to which firms adjust their input choices (Ackerberg et al. 2015).

There are, at least, two important econometric issues in the estimation of equation (5): a selection problem generated by the relationship between the unobserved productivity variable and the shutdown decision, and a simultaneity problem generated by the relationship between productivity and input demand (Olley and Pakes 1996). More specifically, the simultaneity bias arises because the optimal firm's choice of input quantities may depend on the prior beliefs about its productivity level. The existence of such dependence reflects a potential correlation between error term $u_{it}$ and inputs ($k_{it}$ and $l_{it}$), rending OLS estimates of $\beta_k$ and $\beta_l$ inconsistent. The selection bias is likely to occur when only firms with a large capital stock have more chance to stay active. Several methods have been proposed to overcome these problems. Olley and Pakes (1996, henceforth OP) were the first to introduce a semi-parametric estimator that takes both the simultaneity and selection problem explicitly into account. This estimator considers the simultaneity bias by using firm's investment decision to proxy for unobserved productivity shocks. The selection issues, caused by endogenous exit in the sample, are addressed by adding an exit rule into the model. Levinsohn and Petrin (2003, henceforth LP) argue that investment cannot fully respond to productivity shocks. They suggest the use of intermediate inputs such as electricity, fuel oil or materials as proxy variables, the hypothesis being that when productivity increases, production increases, which leads to an increased use of intermediate energy or inputs. More recently, Ackerberg et al. (2015, henceforth ACF) propose an alternative estimation procedure based on the ideas of OP and LP but does not suffer from the functional dependence problems and produces consistent estimates.

Once productivity is estimated, we assume that it is depends on firm-level characteristics ($X_{ijt}$), agglomeration externalities at regional level (localization and urbanization economies) as well as other regional factors such as transportation, human capital, labor concentration ($Z_{jt}$). Formally, the model can be presented as follow:

$$\omega_{ijt} = \gamma_{00} + \gamma_{10}X_{ijt} + \gamma_{01}Z_{jt} + \gamma_{11}X_{ijt}Z_{jt} + [\mu_{0jt} + \mu_{1jt}X_{ijt} + e_{ijt}] \tag{6}$$

where $\omega_{ijt}$ is the productivity of the $i$-th firm (in logarithm) operating in region j at time t (we use the governorate as regional level) and ($X_{ijt}Z_{jt}$) is the interaction term between firm and regional factors. $\gamma_{00}$ is the overall mean across regions or industries. The deterministic part on the model ($\gamma_{00} + \gamma_{10}X_{ijt} + \gamma_{01}Z_{jt} + \gamma_{11}X_{ijt}Z_{jt}$), contains all the fixed coefficients, while the stochastic component is in brackets in equation (6). In order to investigate the differences in productivity between exporter and non-exporter firms and how participation in export markets can influence the firm-level productivity, we added a dummy variable ($\mathrm{export}_{ijt-1}$) among the firm-level characteristics ($X_{ijt}$), that takes a value of 0 if firm i has not exported at time $t-1$, and 1 if it has exported. The model includes also fixed industry effects and time effects. The firm level residuals $e_{ijt}$ are assumed to have a normal distribution with mean zero and variance $\sigma_e^2$. The group-level (regional level) $\mu_{0jt} + \mu_{1jt}$ are assumed to have a multivariate normal distribution with an expected value of zero, and they are assumed to be independent from the individual level residuals $e_{ijt}$ (Van Oort et al 2012). The variances of the residual errors $\mu_{0j}$ and $\mu_{1j}$ are specified as $\sigma^2_{u_0}$ and $\sigma^2_{u_1}$.

For the export model, we use two distinct specifications for the extensive margin and for the intensive margin of exports. In the specification for the extensive margin, we estimate a multilevel logit model that analyses the effects of observable pre-entry firm's characteristics (including productivity in order to capture a potential self-selection process by which more productive firms choose to export) and agglomeration variables at time $(t-1)$ on the probability of exporting at time t ($p_{ijt}$). Formally, the export participation decision for firm i operating in region j at year t ($export_{ijt}$) is expressed by:

$$\log\left(\frac{p_{ijt}}{1-p_{ijt}}\right) = \mathrm{logit}(\mathrm{export}_{ijt}) = \delta_{00} + \delta_{10}X_{ijt-1} + \delta_{01}Z_{jt-1} + \delta_{11}X_{ijt-1}Z_{jt-1} + [u_{0jt} + u_{1jt}X_{ijt-1} + \varepsilon_{ijt}] \quad (7)$$

A similar specification was adopted for the intensive margin of exports, substituting the $\mathrm{export}_{ijt}$ dependent variable with the share of exports over the total sales ($\mathrm{share\_export}_{ijt}$).

$$\mathrm{share\_export}_{ijt} = \delta_{00} + \delta_{10}X_{ijt-1} + \delta_{01}Z_{jt-1} + \delta_{11}X_{ijt-1}Z_{jt-1} + [u_{0jt} + u_{1jt}X_{ijt-1} + \varepsilon_{ijt}] \quad (8)$$

**4. Results**

**4.1. Random intercept models**

We start our analysis by fitting a two-level empty model (without firm-level and regional-level variables) of firm nested within governorates (governorate as level 2 and firm as level 1) for both dependent variables (TFP and export participation decision). The purpose of this step is to test for significant intercept variance, which is a test of the need for mixed modelling. If the intercept variance is not significant, it can be fixed for future steps. For TFP dependent variable, we use the full maximum likelihood (MLF) as well as the restricted or residual maximum likelihood estimation (REML)) to estimate the empty model (see Bryk and Raudenbush 1992 for more details).[8]

The results of MLF and REML are shown in Table 1, where we report the second level intercept ($\gamma_{00}$ for the TFP model and $\delta_{00}$ for the export participation model), its variance and the variance of the lowest level (firm-level). We have also included the intra-class correlation (ICC) and the likelihood ratio test (LR) to compare mixed to non-nested models.

For the TFP model, the LR tests indicate that mixed multilevel models are more appropriate than ordinary linear regressions (the LR tests are significant at the 0.01 level), showing that regional level has a significant impact of firm-level performance and allowing us to justify the use of the multilevel modelling approach. More formally, 3.87% (the value of the intra-class

[8] "Posterior variances will be larger - and more realistic- under REML than under MLF. This will be especially true when the number of level-2 units is small".

correlation) of the difference in firm-level productivity can be explained by regional level (governorates) characteristics, while the remaining difference (96.13%) was attributed to firm-level characteristics. The first graph of Figure 1 plots the random effects across governorates for the Olley-Pakes TFP empty model. We see that the non-coastal governorates, such as Kairouan, Sidi Bouzid, Kasserine and Jendouba have comparatively lower levels of total factor productivity than coastal governorates (such as Ben Arous and Tunis). This gap in TFP between coastal and non-coastal governorates can be explained by at least two factors. First, Tunisia has a considerable infrastructure for production and distribution facilities (major airports, major railways, major roads and agglomeration) concentrated in the coastal area. Second, private capital investments, competitive poles, companies, and jobs are characterized by a regional overconcentration along the coast area (about 90% of total employment and 88% of manufacturing jobs in 2010 are located in coastal areas). Additionally, there is a relative immobility of human capital from the coastal to non-coastal areas. The second two columns of Table 1 report the results of the empty models for firms' export market participation (partially or fully exporting firms (group 1) and fully exporting firms (group 2)). For the first group, the estimating of the intercept provides information that the average probability of exporting is (odds ratio/(1 + odds ratio)) = 82%, while the average probability of exporting for group 2 is only equals to 11%. The computed ICC is equal to 0.342 for the first group and 0.542 for the second group showing that 34.2% and 54.2% of the variation in the probability of exporting can be respectively explained by the regional level (level 2). Those results reveal that there is a significant difference in firms' export market participation across governorates (see figure 2). The last two columns of Table 1 show that 24% of the variation in the share of export can be explained by regional factors.

### 4.2. Random intercept models with only firm-level characteristics

Table 2 shows the results of the random coefficient models with only firm-level characteristics for both dependent variables: total factor productivity and the share of exports over total sales (panel A and panel B, respectively). From left to right, the columns present estimations including an increasing number of controls for firms, industry and year-specific effects. For each panel, column 1 is the baseline estimation with only firm age and firm size variables. It turns out that the majority of explanatory variables are statistically significant at the 5 percent level and they display the expected signs. The results showed that firm size has a positive relationship with firm performance. A positive regression coefficient implies that an increase in firm size (i.e. employ more people) reflects a higher likelihood of firm productivity and firm exporting. Unlike firm size, age does not have the same effect. The coefficient of log age is positive in productivity equation but negative for export equation. The negative relationship between age and exporting can be explained by the fact that older firms are expected to respond less quickly to export orders than younger firms (Williams 2011).

In Tunisia, larger and older firms are more productive, in comparison to smaller and younger firms. Moreover, it is shown that the association between log age and log TFP does not follow a linear pattern (presence of U-shaped relationship, see column 2 in panel A of Table 2). Our results are in line with that part of literature which has reported evidence of positive and significant impacts that size and age have on firm-level productivity (Jovanovic 1982). Our results, however, contradicting those of Marshall (1920), which assume that older firms are unlikely to have the flexibility to make rapid adjustments to changing circumstances and are likely to lose out in performance stakes to younger firms (Majumdar 1997).

The estimated coefficient of the lagged export status of the firm (extensive margin) is positive and statistically significant, which confirms the learning-by-exporting hypothesis and shows that firms sharing contacts with more competitive foreign market tend to be more productive. This result is in line with our expectations and with the finding of Aw et al (2000) for Korea and Van Biesebroeck (2006) for Côte d'Ivoire. However, the coefficient of the lagged share of exports over total sales (intensive margin) is positive and statistically significant only if we control for industry and year effects (last column in panel A of table 2). Ignoring such heterogeneity leads to underestimate the learning by exporting effects, while ignoring the intensive margin could seriously overestimate the learning by exporting effects. The results show also that the hypothesis of self-selection in export markets has been confirmed only after controlling for time and sector-specific effects (the firm-level productivity from the previous year positively affects the export decision, only for the last specification in panel B of Table 2).

**4.3. Random intercept models with firm-level and regional-level characteristics**

Table 3 is complementary to Table 2 and includes variables designed to measure the effect of agglomeration on firm performance. For each dependent variable (firm productivity and firm exporting), we estimated six different models. The first model (column 1) includes only agglomeration economies (specialization and diversity), while the model in column 2 adds the squared values of specialization and diversity to column 1. Model 3 includes both the firm and regional-level variables. The relationship between firm productivity and export behavior is tested by using model 4. Model 5 includes all control variables, except industry and time fixed-effects which are controlled in model 6.

From the results of model 1, it appears that agglomeration economies derive from specialization have a positive effect on firm productivity, but a negative effect on firm exporting. It is however interesting to note that the effect of specialization on firm exporting turns from negative to positive in model 2 (column 2 in panel B of Table 3). The urban density effect, stemming from urbanization economies (diversity) has no impact on firm productivity. However, firm's export volume becomes positively and significantly associated with diversity of the regional economy (last column in panel B). Our results are in line with the finding of Mittelstaedt et al. (2006) who reported that both urbanization and localization economies increase the probability of exporting for small manufacturing firms in the south-Eastern USA in 2002. In the same vein Antonietti and Cainelli (2009) show a positive effect of localization and urbanization economies on both export propensity and export share for Italian manufacturing firms between 1998 and 2003. Other studies, such as Rodríguez-Pose et al. (2012) for Indonesian manufacturing firms, showed that only localization economies have a positive effect on both export propensity and export share.

## 5. Robustness checks and further issues

A series of tests has been performed to check the robustness of the main findings. First, we test if our results are robust to using alternatives approaches (OLS, OP, and ACF) to estimate TFP. Then, we use the quantile approach proposed by Combes et al (2012) in order to distinguish agglomeration from firm selection.

### 5.1. Robustness to alternative measures of TFP

To examine the robustness of our previous findings, for which we used the OP approach to estimate the TFP, we re-estimate our models using the methodology of ACF. Table A.2 in the Appendix reports the descriptive statistics of TFP measures obtained using OLS, OP and ACF as well as the correlations between the different TFP estimators. Mean TFP for the OLS, OP and ACF estimators ranges between 4.52 and 6.97. Both OLS and OP estimators have much higher mean TFP (and also a larger maximum) compared to the ACF estimate (see Figure A.2). Table A.2 shows also that the TFP measures obtained using OLS, OP and ACF are very highly correlated (always higher than 0.91). Table A.3 reports the production function coefficients obtained using the different estimators. Comparing OLS and OP estimates to the ACF estimates in the last row, shows that the coefficients on both log employment and log capital are lower compared to ACF results (the log capital coefficient is not significant for the OP estimate). Both OLS and OP estimates underestimate the effects of input variables on TFP.

To check whether our main results of Table 3 are robust to our choice of TFP estimators, we replicate them using TFP measure obtained from ACF approach. To ease comparisons, the first row of each specification reproduces again the OP estimates presented in Table 3. The next row reports results for the same estimation using the ACF approach. Our results are robust to different estimation approaches. The ACF results maintain the same signs and offer conclusions similar to the ones based on the OP approach. The main differences between the two approaches can be summarized as follows. First, the relationship between export (intensive margin) and productivity becomes positive and statistically significant for the ACF approach (specification 4 in Table 6). The second difference is that the negative effect of firm size becomes statistically significant (specifications 5 and 6 in Table 6).

### 5.2. Agglomeration Economies or Firm Selection?

While most studies in the literature of urban economics have shown that firms and workers are more productive in larger cities (Rosenthal and Strange 2004, Melo et al 2009), some recent studies in this literature have shed new light on another hypothesis of selection (Melitz 2003, Melitz and Ottaviano 2008). Combes et al (2012) nest a generalized version of the firm selection model of Melitz and Ottaviano (2008) and a simple model of agglomeration economies in the spirit of Fujita and Ogawa (1982) and Lucas and Rossi-Hansberg (2002) in order to compare the distribution of firm log productivity across cities of different sizes (larger versus smaller cities). Through this nested model, they were able to parameterize the relative importance of agglomeration and selection and to show that the productivity distributions between larger and smaller cities are comparable via shift, dilation, and truncation components. More specifically, stronger selection effects in larger cities should lead to a greater left truncation of the productivity distribution by excluding the least productive firms (due to tougher competition). Meanwhile, stronger agglomeration effects are expressed by shifting productivity distribution of larger cities rightward (all firms in larger cities enjoy the same benefits from agglomeration) and by making it dilated (more productive firms in larger cities enjoy greater benefits from agglomeration) (Kondo 2017).

Since it is hard to separate truncation, shift, and dilation in a purely visual comparison of distributions, Combes et al (2012) propose a new quantile approach to distinguish agglomeration effects from firm selection. This approach estimates the extent to which the log productivity distribution in larger cities is left-truncation or dilated and right-shifted compared to the log productivity distribution in smaller cities. Formally, Combes et al (2012) estimate three parameters: $A$, $D$ and $S$. The parameter $A$ measures how much stronger the right shift of log productivity distribution in larger cities relative to smaller cities; parameter $D$ measures the ratio of dilation in larger cities relative to smaller cities; and parameter $S$ measures how much stronger the left truncation of the log productivity distribution in larger cities relative to smaller cities (see Combes et al 2012 for more details). If $A > 0$ and values of $D$ above unity, more productive firms benefits more from being in larger cities (in terms of employment density), whereas if values of $D$ are below unity, more productive firms benefit less from being in larger cities. Positive values of $S$ are evidence that the distribution of firm log productivity in larger cities is more truncated than in smaller cities (selection effect).

Columns 1, 2 and 3 of Table 7 report the estimation results of $A$, $D$ and $S$ with bootstrapped standard errors for the three different approach (OLS, OP and ACF) and for five different specifications (Shift; Shift + Dilation; Shift + Truncation; Truncation; and Shift + Dilation + Truncation). We consider as bigger cities the three agglomerations of Great Tunis, Sousse and Sfax (the three main cities in Tunisia) and all other cities are considered as smaller cities (three biggest versus other cities). All values of $A$ are positive and significant at 1 percent level, and values of $D$ are below unity which means that more productive manufacturing firms benefit less from being in one of the three biggest cities in Tunisia. Those results confirm the previous finding of empty model which show that only 3.87% of the difference in firm-level productivity can be explained by regional level variables (approximated here by agglomeration economies). In addition, values of $S$ are positive and significant at 1 percent level showing that the distribution of firm log productivity in larger cities is more truncated than in smaller cities (selection effect, allowing only the most productive firms to survive).

## 6. Conclusion and policy relevance

This paper assessed the role of agglomeration effects on total factor productivity and the export decision of manufacturing firms between 1998 and 2004 for Tunisia. Our results support for both self-selection and learning by exporting hypotheses. This is in line with the findings by Bernard and Wagner (2001) supporting the self-selection hypothesis and are also in line with Greenaway and Kneller (2004) supporting the learning-by-exporting hypothesis. The results of the multilevel analysis, which considers firms as the first level group in the hierarchy of data and governorates as the second-level group, show that the governorate contributes almost 3.87% to total factor productivity. These results imply that the unobserved heterogeneity in firm behavior is the main source of heterogeneity in productivity. In this respect, the government should tailor micro policies such as managerial support/guidance, innovation and human capital rather than public investments.

The results of multilevel analysis show, however, that regional characteristics boost export capacity of manufacturing firms (the governorate contributes about 34.2% of the export participation decision). This is why policy-makers should be interest to both micro and regional level factors, where cities are still unable to capitalize on the growth generated by agglomeration economies to further boost export capacity. For example, public investments (such as road infrastructure, industrial parks, etc.) have taken place in a context where the government lack adequately designed urban policies to amplify the benefits of agglomeration. It has thus become clear that urban investments need to be better targeted and their quality enhanced. As such, most policy-makers are unable to strategically use public investments to attract investments which generate opportunities for sustainable development and growth at the local and national levels. They often require more than just a general sense that urban agglomeration economies exist. Instead, they would presumably be interested in whether agglomeration economies are stronger and more beneficial in some areas than others in order to tailor spatially targeted interventions to help places that are not doing well and to stimulate the others that are already doing well to continue to do better.

**References**


Ackerberg, D. A., Caves, K., & Frazer, G. (2015) Identification properties of recent production function estimators. *Econometrica*, 83(6), 2411-2451.

Aiello, F., & Ricotta, F. (2016) Firm heterogeneity in productivity across Europe: evidence from multilevel models. *Economics of Innovation and New Technology*, 25(1), 57-89.

Aiello, F., Pupo, V., & Ricotta, F. (2014) Explaining total factor productivity at firm level in Italy: Does location matter? *Spatial Economic Analysis*, 9(1), 51-70.

Alker, H.R. (1969) A typology of ecological fallacies. In M. Dogan and S. Rokan (eds), *Quantitative Ecological Analysis* (pp. 69-86). Cambridge, MA: MIT Press.

Amara, M., & Ayadi, M. (2014) Local employment growth in the coastal area of Tunisia: spatial filtering approach. *Middle East Development Journal*, 6(2), 255-284.

Amara, M., & Thabet, K. (2016) Firm and Regional Factors of Productivity: A Multilevel Analysis of Tunisian Manufacturing. *Economic Research Forum* Working Papers (No. 1041).

Antonietti, R., & Cainelli, G. (2011) The role of spatial agglomeration in a structural model of innovation, productivity and export: a firm-level analysis. *The Annals of Regional Science*, 46(3), 577-600.

Arrow, K. (1962) The Economic Implications of Learning by Doing. *The Review of Economic Studies*, 29(3), 155-173.

Audretsch, D. B., & Feldman, M. P. (1996) R&D spillovers and the geography of innovation and production. *The American economic review*, 86(3), 630-640.

Audretsch, D. B., & Feldman, M. P. (2004) *Knowledge spillovers and the geography of innovation*. In Handbook of regional and urban economics, 4, 2713-2739.

Aw, B. Y., Chung, S., & Roberts, M. J. (2000) Productivity and turnover in the export market: micro-level evidence from the Republic of Korea and Taiwan (China). *The World Bank Economic Review*, 14(1), 65-90.

Ayyagari, M., Demirgüç-Kunt, A., & Maksimovic, V. (2011) Small vs. young firms across the world: contribution to employment, job creation, and growth, Policy Research Working Paper No. 5631. World Bank, Washington, DC.

Baccouche, R., Bouoiyour, J., M'Henni, H., & Mouley, S. (2008) Dynamics of investments, changing industry and convertibility capital account: impacts of the measures of liberalization and compared experiences Tunisia-Morocco, Research Project N° FEM 32- 04.

Backman, M. (2014) Human capital in firms and regions: Impact on firm productivity. *Papers in Regional Science*, 93(3), 557-575.

Ballot, G., FakhFakh, F., & Taymaz, E. (2001) Firms' human capital, R&D and performance: a study on French and Swedish firms. *Labour economics*, 8(4), 443-462.

Beaudry, C., & Schiffauerova, A. (2009) Who's right, Marshall or Jacobs? The localization versus urbanization debate. *Research policy*, 38(2), 318-337.

Bernard, A. B., & Jensen, J. B. (1999) Exceptional exporter performance: cause, effect, or both? *Journal of international economics*, 47(1), 1-25.

Bernard, A. B., & Jensen, J. B. (2004) Why some firms export. *Review of economics and Statistics*, 86(2), 561-569.

Bronzini, R., & Piselli, P. (2009) Determinants of long-run regional productivity with geographical spillovers: the role of R&D, human capital and public infrastructure. *Regional Science and Urban Economics*, 39(2), 187-199.

Bryk, A. S., & Raudenbush, S. W. (1992) *Advanced qualitative techniques in the social sciences, 1. Hierarchical linear models: Applications and data analysis methods.* Thousand Oaks, CA, US: Sage Publications, Inc.

Cainelli, G., Ganau, R., & Iacobucci, D. (2016) Do Geographic Concentration and Vertically Related Variety Foster Firm Productivity? Micro-Evidence from I taly. *Growth and Change*, 47(2), 197-217.

Ciccone, A., Hall, R.E. (1996) Productivity and the density of economic activity. *American Economic Review*, 86, 54-70.

Cingano, F., & Schivardi, F. (2004) Identifying the sources of local productivity growth. *Journal of the European Economic association*, 2(4), 720-742.

Clerides, S. K., Lach, S., & Tybout, J. R. (1998) Is learning by exporting important? Micro-dynamic evidence from Colombia, Mexico, and Morocco. *The Quarterly Journal of Economics*, 113(3), 903-947.

Cohen, J. P., & Paul, C. J. M. (2005) Agglomeration economies and industry location decisions: the impacts of spatial and industrial spillovers. *Regional Science and Urban Economics*, 35(3), 215-237.

Cohen, J. P., & Paul, C. J. M. (2009) Agglomeration, productivity and regional growth: production theory approaches. *Handbook of regional growth and development theories*, 101-117.

Combes, P. P. (2000) Economic structure and local growth: France, 1984-1993. *Journal of urban economics*, 47(3), 329-355.

Combes, P. P., Duranton, G., Gobillon, L., Puga, D., & Roux, S. (2012) The productivity advantages of large cities: Distinguishing agglomeration from firm selection. *Econometrica*, 80(6), 2543-2594.

Combes, P. P., Magnac, T., & Robin, J. M. (2004). The dynamics of local employment in France. *Journal of Urban Economics*, 56(2), 217-243

De, P. K., & Nagaraj, P. (2014) Productivity and firm size in India. *Small Business Economics*, 42(4), 891-907.

Diop, N. (2008) *Tunisia's global integration: A second generation of reforms to boost growth and employment*. Washington, DC: World Bank, 1-145

Duranton, G., & Puga, D. (2000) Diversity and specialisation in cities: why, where and when does it matter?. *Urban studies*, 37(3), 533 -555.

Duranton, G., & Puga, D. (2004) *Micro-foundations of urban agglomeration economies*. In Handbook of regional and urban economics, 4, 2063-2117.

Ellison, G., & Glaeser, E. L. (1997) Geographic concentration in US manufacturing industries: a dartboard approach. *Journal of political economy*, 105(5), 889-927.

Fazio, G., & Piacentino, D. (2010) A Spatial Multilevel Analysis of Italian SMEs' Productivity. *Spatial Economic Analysis*, 5(3), 299-316.

Fracasso, A., & Vittucci Marzetti, G. (2018) Estimating dynamic localization economies: the inadvertent success of the specialization index and the location quotient. *Regional Studies*, 52(1), 119-132.

Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., & Shleifer, A. (1992) Growth in cities. *Journal of political economy*, 100(6), 1126-1152.

Harris, R., & Li, Q. C. (2008) Exporting, R&D, and absorptive capacity in UK establishments. *Oxford economic papers*, 61(1), 74-103.

Henderson, J. V. (2007) Understanding knowledge spillovers. *Regional Science and Urban Economics*, 37(4), 497-508.

Henderson, J. V., & Wang, H. G. (2007) Urbanization and city growth: The role of institutions. *Regional Science and Urban Economics*, 37(3), 283-313.

Henderson, V. (2003) The urbanization process and economic growth: The so-what question. *Journal of Economic growth*, 8(1), 47-71.

Henderson, V., Kuncoro, A., & Turner, M. (1995) Industrial development in cities. *Journal of Political Economy*, 103(5), 1067-1090.

Jacobs, J. (1969) *The economy of cities*. Random House, New York.

Jelili, R. B., & Goaied, M. (2010) Entry, exit, and productivity in tunisian manufacturing industries. In *Market Dynamics and Productivity in Developing Countries* (pp. 73-108). Springer, New York, NY.

Jones, J. (2017) Agglomeration economies and the location of foreign direct investment: A meta-analysis. *Journal of Regional Science*, 57(5), 731-757.

Jovanovic, B. (1982). Selection and the Evolution of Industry. *Econometrica,* 50(3), 649-670.

Ke, S. (2010) Agglomeration, productivity, and spatial spillovers across Chinese cities. *The Annals of Regional Science*, 45(1), 157-179.

Keller, W., & Yeaple, S. R. (2009) Multinational enterprises, international trade, and productivity growth: firm-level evidence from the United States. *The Review of Economics and Statistics*, 91(4), 821-831.

Kondo, K (2017) Quantile Approach for Distinguishing Agglomeration from Firm Selection in Stata, RIETI Technical Paper 17-T-001.

Kox, H. L., & Rojas-Romagosa, H. (2010) Exports and productivity selection effects for Dutch firms. *De Economist*, 158(3), 295-322.

Krugman, P. R. (1991) *Geography and trade*. MIT press.

Levinsohn, J., & Petrin, A. (2003) Estimating Production Functions Using Inputs to Control for Unobservables, *Review of Economic Studies*, 70 (2), 317-342.

Majumdar, S. K. (1997) The impact of size and age on firm-level performance: some evidence from India. *Review of industrial organization*, 12(2), 231-241.

Marouani, M. A., & Mouelhi, R. (2016) Contribution of Structural Change to Productivity Growth: Evidence from Tunisia. *Journal of African Economies*, 25(1): 110-132.

Marshall, A. (1920) *Principles of economics*, Macmillan and Co., London.

Martin, P., Mayer, T., & Mayneris, F. (2011) Spatial concentration and plant-level productivity in France. *Journal of Urban Economics*, 69(2), 182-195.

Melo, P. C., Graham, D. J., & Noland, R. B. (2009) A meta-analysis of estimates of urban agglomeration economies. *Regional science and urban Economics*, 39(3), 332-342.

Mittelstaedt, J. D., Ward, W. A., & Nowlin, E. (2006) Location, industrial concentration and the propensity of small US firms to export: Entrepreneurship in the international marketplace. *International Marketing Review*, 23(5), 486-503.

Moretti, E. (2004) Workers' education, spillovers, and productivity: evidence from plant level production functions. *The American Economic Review*, 94(3), 656-690.

Neffke, F., Henning, M., Boschma, R., Lundquist, K. J., & Olander, L. O. (2011) The dynamics of agglomeration externalities along the life cycle of industries. *Regional studies*, 45(1), 49-65.

Parisi, M. L., Schiantarelli, F., & Sembenelli, A. (2006) Productivity, innovation and R&D: Micro evidence for Italy. *European Economic Review*, 50(8), 2037-2061.

Raspe, O., & van Oort, F. (2011a) Growth of new firms and spatially bounded knowledge externalities. *The annals of regional science*, 46(3), 495-518.

Raspe, O., & van Oort, F. G. (2011b) Firm heterogeneity, productivity and spatially bounded knowledge externalities. *Socio-Economic Planning Sciences*, 45(1), 38-47.

Rice, P., Venables, A. J., & Patacchini, E. (2006) Spatial determinants of productivity: analysis for the regions of Great Britain. *Regional Science and Urban Economics*, 36(6), 727-752.

Roberts, M. J., & Tybout, J. R. (1997) The decision to export in Colombia: an empirical model of entry with sunk costs. *The American Economic Review*, 87(4), 545-564.

Robinson, W.S. (1950) Ecological correlations and the behavior of individuals. *Sociological Review,* 15, 351-357.

Rodríguez-Pose, A., & Hardy, D. (2017) Firm competitiveness and regional disparities in Georgia. *Geographical Review*, 107(2), 384-411.

Romer, P. (1990) Endogenous technological change. Technical report.

Rosenthal, S. S., & Strange, W. C. (2004) Evidence on the nature and sources of agglomeration economies. In Handbook of regional and urban economics (Vol. 4, pp. 2119-2171). Elsevier

Snijders, T. A. B. & Bosker, R. J. (1999) *Multilevel Analysis: an Introduction to Basic and Advanced Multilevel Modeling*. London: Sage

Van Biesebroeck, J. (2005) Exporting raises productivity in sub-Saharan African manufacturing firms. *Journal of International Economics*, 67(2), 373-391.

van Dijk, J., Folmer, H., & Oosterhaven, J. (2009) *Regional policy: rationale, foundations and measurement of its effects*. In Handbook of regional growth and development theories, 461.

Van Oort, F. G., Burger, M. J., Knoben, J., & Raspe, O. (2012) Multilevel Approaches and the Firm-Agglomeration Ambiguity in Economic Growth Studies. *Journal of Economic Surveys*, 26(3), 468-491

Wagner, J. (2007) Exports and productivity: A survey of the evidence from firm-level data. *The World Economy*, 30(1), 60-82.

Wakelin, K. (2001) Productivity growth and R&D expenditure in UK manufacturing firms. *Research policy*, 30(7), 1079-1090.

Williams, D. A. (2011) Impact of firm size and age on the export behaviour of small locally owned firms: Fresh insights. *Journal of International Entrepreneurship*, 9(2), 152-174.

World Bank. (2009) *World development report 2009: Reshaping economic geography*. Washington, DC: The World Bank.

World Bank. (2014) *The unfinished revolution: bringing opportunity, good jobs and greater wealth to all Tunisians*. Washington, DC: World Bank Group.

**Table 1: Empty model for total factor productivity, export participation decision and the share of exports over total sales**

| | Total Factor Productivity (OP) | | Export participation Decision | | Share of exports over total sales | |
|---|---|---|---|---|---|---|
| | MLF | REML | fully or partially exporting firms | Fully exporting firms | MLF | REML |
| Constant/odds ratio | 6.887*** | 6.918*** | 4.878*** | 0.125*** | 0.405*** | 0.404*** |
| standard error | 0.033 | 0.038 | 1.369 | 0.053 | 0.057 | 0.059 |
| Variance of the error term at level 2 (governorate) | 0.012*** | 0.022*** | 1.711*** | 3.888*** | 0.056*** | 0.060*** |
| Variance of the error term at level 1 (firm) | 0.544*** | 0.542*** | $\pi^2/3 \approx 3.29$ | $\pi^2/3 \approx 3.29$ | 0.187*** | 0.187*** |
| ICC | 2.14% | 3.87% | 34.2% | 54.2% | 23.06% | 24.22% |
| LR chi(2) | 186.27*** | 211.48*** | 865*** | 21467*** | 1448.64*** | 1453*** |
| Log likelihhod | -9084 | | -5522 | -4989 | -4736 | |
| Log restricted-likelihood | | -9075 | | | | -4738 |

*Notes*: *** significant at 1% level. The ICC is the ratio between the variance of level 2 and the total variance (variance of level 2 + variance of level 1).

**Table 2: Random intercept TFP models with only firm-level characteristics**

| Dependent variables | *Panel A: Total Factor Productivity* (OP) | | | | | | *Panel B : Share of exports over total sales* | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (1) | (2) | (3) | (4) | (5) |
| *Firm level variables* | | | | | | | | | | | |
| Log of size | 0.161*** | -0.062 | -0.145** | -0.111** | -0.020 | 0.014 | 0.166*** | 0.264*** | 0.256*** | 0.272*** | 0.080*** |
| | (0.008) | (0.043) | (0.051) | (0.051) | (0.049) | (0.049) | (0.004) | (0.023) | (0.028) | (0.028) | (0.025) |
| Log of age | 0.019* | -0.156*** | -0.462*** | -0.459*** | -0.385*** | -0.385*** | -0.187*** | -0.008 | 0.035 | 0.026 | 0.007 |
| | (0.010) | (0.041) | (0.070) | (0.070) | (0.067) | (0.068) | (0.005) | (0.021) | (0.040) | (0.040) | (0.034) |
| Log of size squared | | 0.026*** | 0.033*** | 0.031*** | 0.022*** | 0.018*** | | -0.011*** | -0.010*** | -0.012*** | 0.002 |
| | | (0.005) | (0.006) | (0.006) | (0.005) | (0.005) | | (0.003) | (0.003) | (0.003) | (0.003) |
| Log of age squared | | 0.038*** | 0.096*** | 0.093*** | 0.062*** | 0.066*** | | -0.039*** | -0.048*** | -0.046*** | -0.027*** |
| | | (0.009) | (0.014) | (0.014) | (0.013) | (0.013) | | (0.005) | (0.008) | (0.008) | (0.007) |
| Fully or partially exporting | | | 0.077*** | | 0.185*** | | | | | | |
| at (t-1) (extensive margin) | | | (0.022) | | (0.022) | | | | | | |
| Share of exports over total sales | | | | -0.020 | | 0.211*** | | | | | |
| at (t-1) (intensive margin) | | | | (0.024) | | (0.027) | | | | | |
| Total factor productivity | | | | | | | | | -0.006 | -0.001 | 0.042*** |
| at (t-1) (OP approach) | | | | | | | | | (0.008) | (0.008) | (0.009) |
| Dummy R&D | | | | | 0.112*** | 0.122*** | | | | -0.073*** | -0.040*** |
| | | | | | (0.018) | (0.019) | | | | (0.011) | (0.009) |
| Constant | 6.175*** | 6.791*** | 7.453*** | 7.410*** | 7.265*** | 7.186*** | 0.156*** | -0.214** | -0.228** | -0.270** | -0.323*** |
| | (0.055) | (0.108) | (0.147) | (0.147) | (0.162) | (0.163) | (0.046) | (0.068) | (0.106) | (0.106) | (0.096) |
| Observations | 7805 | 7805 | 5274 | 5250 | 5274 | 5250 | 7773 | 7773 | 5260 | 5260 | 5260 |
| Number of governorates | 23 | 23 | 22 | 22 | 22 | 22 | 23 | 23 | 22 | 22 | 22 |
| Sector fixed effects | No | No | No | No | Yes | Yes | No | No | No | No | Yes |
| Time fixed effects | No | No | No | No | Yes | Yes | No | No | No | No | Yes |
| ICC | 4.57% | 4.62% | 4.56% | 4.34% | 3.92% | 4.54% | 19.18% | 20.02% | 21.42% | 21.74% | 9.66% |
| LR test vs. linear regression | 173.38*** | 159.96*** | 125.44*** | 112.68*** | 76.37*** | 86.27*** | 683.17*** | 642.85*** | 415.38*** | 369.12*** | 223.12*** |
| Log likelihood | -8441 | -8426 | -5343 | -5327 | -5047 | -5028 | -3477 | -3440 | -2321 | -2303 | -1521 |

Standard errors in parentheses. *** significant at 1%; ** significant at 5%; * significant at 10%. For the fist and the second specifications the number of governorates is equal to 23 because the governorate of Kebili does not present any firm. For specifications with 22 governorates, the governorate of Kebili and Tozeur do not present any firm.

## Table 3: Random intercept TFP models with firm-level and regional-level characteristic

| | *Panel A: Total Factor Productivity* (OP) | | | | | | *Panel B : Share of exports over total sales* | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (1) | (2) | (3) | (4) | (5) | (6) |
| ***Firm level variables*** | | | | | | | | | | | | |
| Log of size | | | -0.173*** | -0.139** | -0.021 | 0.015 | | | 0.261*** | 0.254*** | 0.270*** | 0.081*** |
| | | | (0.051) | (0.051) | (0.049) | (0.049) | | | (0.023) | (0.028) | (0.028) | (0.025) |
| Log of age | | | -0.431*** | -0.428*** | -0.381*** | -0.382*** | | | -0.008 | 0.041 | 0.032 | 0.010 |
| | | | (0.069) | (0.070) | (0.067) | (0.067) | | | (0.021) | (0.040) | (0.039) | (0.034) |
| Log of size squared | | | 0.034*** | 0.032*** | 0.021*** | 0.018*** | | | -0.011*** | -0.010*** | -0.012*** | 0.001 |
| | | | (0.006) | (0.006) | (0.005) | (0.005) | | | (0.003) | (0.003) | (0.003) | (0.003) |
| Log of age squared | | | 0.087*** | 0.084*** | 0.061*** | 0.065*** | | | -0.039*** | -0.049*** | -0.046*** | -0.027*** |
| | | | (0.014) | (0.014) | (0.013) | (0.013) | | | (0.005) | (0.008) | (0.008) | (0.007) |
| Fully or partially exporting | | | 0.089*** | | 0.184*** | | | | | | | |
| at (t-1) (extensive margin) | | | (0.021) | | (0.022) | | | | | | | |
| Share of exports over total | | | | 0.003 | | 0.201*** | | | | | | |
| sales at (t-1) (intensive margin) | | | | (0.024) | | (0.027) | | | | | | |
| Total factor productivity | | | | | | | | | | -0.005 | 0.0001 | 0.048*** |
| at (t-1) (OP approach) | | | | | | | | | | (0.008) | (0.008) | (0.007) |
| Dummy R&D | | | 0.152*** | 0.152*** | 0.111*** | 0.121*** | | | | | -0.071*** | -0.039*** |
| | | | (0.019) | (0.019) | (0.018) | (0.019) | | | | | (0.011) | (0.009) |
| ***Regional level variables*** | | | | | | | | | | | | |
| Specialisation | 0.170*** | 0.254*** | 0.214*** | 0.219*** | 0.126*** | 0.113** | -0.033*** | 0.104*** | 0.081*** | 0.075*** | 0.075*** | 0.073*** |
| | (0.016) | (0.036) | (0.040) | (0.040) | (0.038) | (0.039) | (0.009) | (0.022) | (0.019) | (0.023) | (0.023) | (0.020) |
| Relative diversity | -0.004 | 0.104 | -0.143 | -0.133 | 0.0893 | 0.082 | 0.015 | 0.006 | 0.051 | 0.091 | 0.089 | 0.122** |
| | (0.017) | (0.070) | (0.092) | (0.092) | (0.077) | (0.080) | (0.011) | (0.055) | (0.044) | (0.057) | (0.057) | (0.048) |
| Specialisation squared | | -0.025*** | -0.021** | -0.023*** | -0.013 | -0.012 | | -0.041*** | -0.032*** | -0.030*** | -0.030*** | -0.011** |
| | | (0.009) | (0.011) | (0.011) | (0.010) | (0.010) | | (0.006) | (0.005) | (0.006) | (0.006) | (0.005) |
| Relative diversity squared | | -0.012 | 0.018 | 0.017 | -0.009 | -0.008 | | 0.0003 | -0.004 | -0.010 | -0.010 | -0.014** |
| | | (0.008) | (0.011) | (0.011) | (0.010) | (0.010) | | (0.006) | (0.005) | (0.007) | (0.007) | (0.006) |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Constant | 6.627*** | 6.449*** | 7.424*** | 7.364*** | 6.924*** | 6.881*** | 0.424*** | 0.406*** | -0.287** | -0.357*** | -0.397*** | -0.591*** |
| | (0.076) | (0.127) | (0.2157) | (0.216) | (0.209) | (0.212) | (0.054) | (0.100) | (0.096) | (0.137) | (0.137) | (0.123) |
| Observations | 8139 | 8139 | 5274 | 5250 | 5274 | 5250 | 8106 | 8106 | 7773 | 5260 | 5260 | 5260 |
| Number of governorates | 23 | 23 | 22 | 22 | 22 | 22 | 23 | 23 | 22 | 22 | 22 | 22 |
| Sector fixed effects | No | No | No | No | Yes | Yes | No | No | No | No | No | Yes |
| Time fixed effects | No | No | No | No | Yes | Yes | No | No | No | No | No | Yes |
| ICC | 9.92% | 4.13% | 7.61% | 7.42% | 3.63% | 4.27% | 17.74% | 15.33% | 10.10% | 11.83% | 11.56% | 8.71% |
| LR test vs. linear regression | 220.15*** | 183.77*** | 110.89*** | 105.59*** | 67.39*** | 76.35*** | 1357*** | 1096*** | 469*** | 279.60*** | 250.58*** | 217.74*** |
| Log likelihood | -9024 | -9027 | -5289 | -5275 | -5049 | -5033 | -4719 | -4703 | -3428 | -2316 | -2298 | -1524 |

Standard errors in parentheses. *** significant at 1%; ** significant at 5%; * significant at 10%. For the fist and the second specifications the number of governorates is equal to 23 because the governorate of Kebili does not present any firm. For specifications with 22 governorates, the governorate of Kebili and Tozeur do not present any firm.

**Table 4: Random intercept TFP models with firm-level and regional-level characteristics: Robustness to alternative estimation approaches (OP vs. ACF)**

| | (1) | | (2) | | (3) | | (4) | | (4) | | (5) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (OP) | (ACF) | (OP) | (ACF) | (OP) | (ACF) | (OP) | (ACF) | (OP) | (ACF) | (OP) | (ACF) |
| ***Firm level variables*** | | | | | | | | | | | | |
| Log of size | | | | | -0.173*** | -0.254*** | -0.139** | -0.259*** | -0.021 | -0.176*** | 0.015 | -0.144** |
| | | | | | (0.051) | (0.048) | (0.051) | (0.048) | (0.049) | (0.045) | (0.049) | (0.047) |
| Log of age | | | | | -0.431*** | -0.507*** | -0.428*** | -0.502*** | -0.381*** | -0.493*** | -0.382*** | -0.497*** |
| | | | | | (0.069) | (0.066) | (0.070) | (0.066) | (0.067) | (0.065) | (0.067) | (0.065) |
| Log of size squared | | | | | 0.034*** | 0.022*** | 0.032*** | 0.022*** | 0.021*** | 0.015** | 0.018*** | 0.011** |
| | | | | | (0.006) | (0.005) | (0.006) | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) |
| Log of age squared | | | | | 0.087*** | 0.088*** | 0.084*** | 0.087*** | 0.061*** | 0.072*** | 0.065*** | 0.079*** |
| | | | | | (0.014) | (0.013) | (0.014) | (0.013) | (0.013) | (0.013) | (0.013) | (0.013) |
| Fully or partially exporting at (t-1) (extensive margin) | | | | | 0.089*** | 0.188*** | | | 0.184*** | 0.209*** | | |
| | | | | | (0.021) | (0.022) | | | (0.022) | (0.021) | | |
| Share of exports over total sales at (t-1) (intensive margin) | | | | | | | 0.003 | 0.212*** | | | 0.201*** | 0.321*** |
| | | | | | | | (0.024) | (0.023) | | | (0.027) | (0.026) |
| Dummy R&D | | | | | 0.152*** | 0.117*** | 0.152*** | 0.117*** | 0.111*** | 0.081*** | 0.121*** | 0.094*** |
| | | | | | (0.019) | (0.018) | (0.019) | (0.018) | (0.018) | (0.018) | (0.019) | (0.018) |
| ***Regional level variables*** | | | | | | | | | | | | |
| Specialisation | 0.170*** | 0.125*** | 0.254*** | 0.214*** | 0.214*** | 0.224*** | 0.219*** | 0.222*** | 0.126*** | 0.164*** | 0.113** | 0.137*** |
| | (0.016) | (0.015) | (0.036) | (0.034) | (0.040) | (0.037) | (0.040) | (0.037) | (0.038) | (0.037) | (0.039) | (0.037) |
| Relative diversity | -0.004 | 0.003 | 0.104 | 0.096 | -0.143 | -0.068 | -0.133 | -0.068 | 0.089 | 0.085 | 0.082 | 0.071 |
| | (0.017) | (0.016) | (0.070) | (0.065) | (0.092) | (0.079) | (0.092) | (0.079) | (0.077) | (0.067) | (0.080) | (0.072) |
| Specialisation Squared | | | -0.025*** | -0.027** | -0.021** | -0.027** | -0.023*** | -0.026** | -0.013 | -0.020** | -0.012 | -0.016* |
| | | | (0.009) | (0.009) | (0.011) | (0.010) | (0.011) | (0.010) | (0.010) | (0.009) | (0.010) | (0.009) |
| Relative diversity Squared | | | -0.012 | -0.010 | 0.018 | 0.011 | 0.017 | 0.011 | -0.009 | -0.008 | -0.008 | -0.006 |
| | | | (0.008) | (0.007) | (0.011) | (0.010) | (0.011) | (0.010) | (0.010) | (0.009) | (0.010) | (0.009) |
| Constant | 6.627*** | 4.186*** | 6.449*** | 4.035*** | 7.424*** | 5.571*** | 7.364*** | 5.568*** | 6.924*** | 5.347*** | 6.881*** | 5.320*** |
| | (0.076) | (0.076) | (0.127) | (0.119) | (0.2157) | (0.194) | (0.216) | (0.194) | (0.209) | (0.194) | (0.212) | (0.199) |
| Observations | 8139 | 8139 | 8139 | 8139 | 5274 | 5274 | 5250 | 5250 | 5274 | 5274 | 5250 | 5250 |
| Number of governorates | 23 | 23 | 23 | 23 | 22 | 22 | 22 | 22 | 22 | 22 | 22 | 22 |
| Sector fixed effects | No | No | No | No | No | No | No | No | Yes | Yes | Yes | Yes |
| Time fixed effects | No | No | No | No | No | No | No | No | Yes | Yes | Yes | Yes |
| ICC | 9.92% | 11.73% | 4.13% | 3.95% | 7.61% | 4.62% | 7.42% | 4.49% | 3.63% | 1.95% | 4.27% | 2.96% |
| LR test vs. linear regression | 220.15*** | 67.44*** | 183.77*** | 32.07*** | 110.89*** | 51.05*** | 105.59*** | 48.16*** | 67.39*** | 26.05*** | 76.35*** | 37.70*** |
| Log likelihood | -9024 | -8551 | -9027 | -8555 | -5289 | -5031 | -5275 | -5002 | -5049 | -4889 | -5033 | -4839 |

*Notes:* Standard errors in parentheses. *** significant at 1%; ** significant at 5%; * significant at 10%. For the fist and the second specifications the number of governorates is equal to 23 because the governorate of Kebili does not present any firm. For specifications with 22 governorates, the governorate of Kebili and Tozeur do not present any firm.

**Table 5: Agglomeration or firm selection: robustness to alternative estimation approaches (OLS, OP, ACF)**

| Model specification | $\hat{A}$ | $\hat{D}$ | $\hat{S}$ | Pseudo-$R^2$ | Obs |
|---|---|---|---|---|---|
| **Shift** | | | | | |
| OLS | 0.098*** (0.019) | | | 0.146 | 8017 |
| OP | 0.122*** (0.020) | | | 0.234 | 8017 |
| ACF | 0.078*** (0.014) | | | 0.487 | 8017 |
| **Shift + Dilation** | | | | | |
| OLS | 0.078*** (0.014) | 0.797*** (0.036) | | 0.487 | 8017 |
| OP | 0.104*** (0.016) | 0.850*** (0.034) | | 0.444 | 8017 |
| ACF | 0.078*** (0.014) | 0.793*** (0.036) | | 0.498 | 8017 |
| **Shift + Truncation** | | | | | |
| OLS | 0.049*** (0.018) | | 0.017*** (0.006) | 0.951 | 8017 |
| OP | 0.077*** (0.020) | | 0.016*** (0.005) | 0.980 | 8017 |
| ACF | 0.048*** (0.018) | | 0.018*** (0.006) | 0.944 | 8017 |
| **Truncation** | | | | | |
| OLS | | | 0.028*** (0.007) | 0.925 | 8017 |
| OP | | | 0.031*** (0.007) | 0.913 | 8017 |
| ACF | | | 0.028*** (0.007) | 0.920 | 8017 |
| **Shift + dilation + Truncation** | | | | | |
| OLS | 0.058*** (0.015) | 0.920*** (0.026) | 0.010*** (0.004) | 0.974 | 8017 |
| OP | 0.080*** (0.018) | 0.978 (0.026) | 0.013*** (0.005) | 0.982 | 8017 |
| ACF | 0.059*** (0.015) | 0.910*** (0.026) | 0.009** (0.004) | 0.973 | 8017 |

*Notes*: ***, **, * For $\hat{A}$ and $\hat{S}$ significantly different from 0 at 1%, 5% and 10%; for $\hat{D}$ significanlty different from 1 at 1%, 5% and 10%, respectively. The bootstrap standard errors are in parentheses. The pseudo-$R^2 = 1 - \frac{M(\hat{A},\hat{D},\hat{S})}{M(0,1,0)}$, $M$ is the criteria function to be minimized, which is defined as the sum of the squared values obtained from the empirical quantile functions of larger and smaller cities (see Combes et al. 2012 for the more details).The pseudo-$R^2$ becomes 0 if the two distributions are identical ($\hat{A}$ = 0, $\hat{D}$ = 1 and $\hat{S}$ = 0).

**Figure 1: Random effects of the intercept for the log TFP**

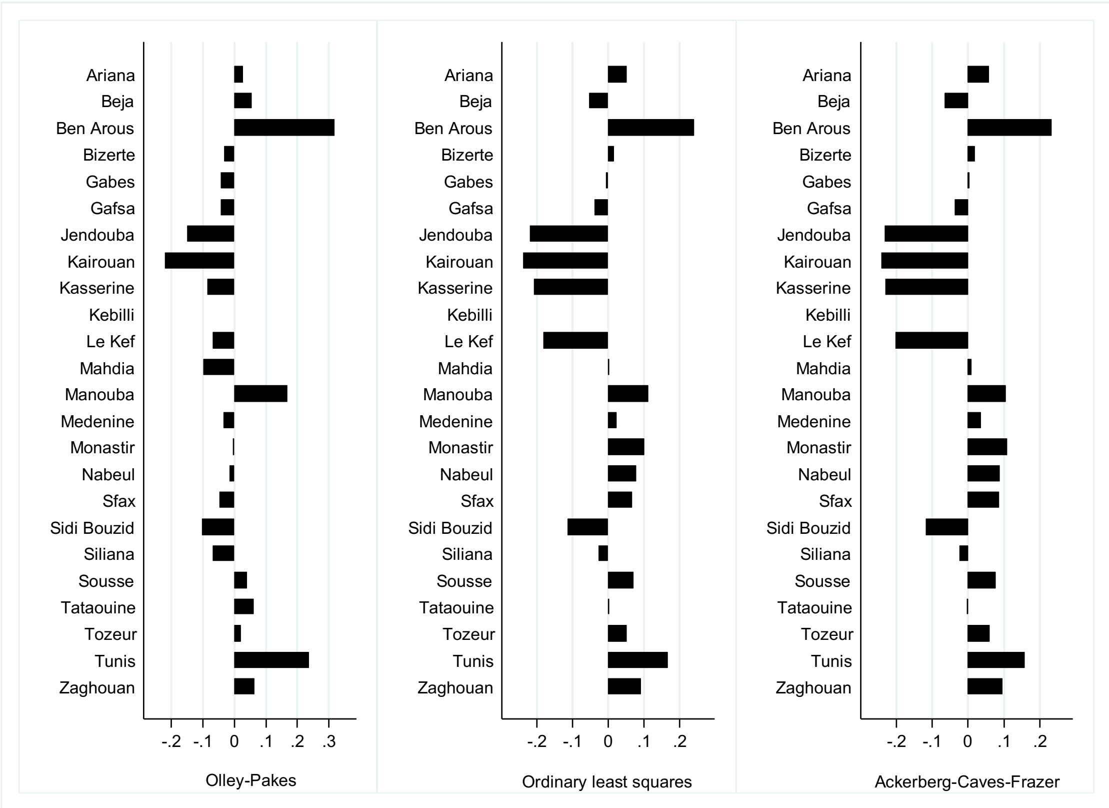
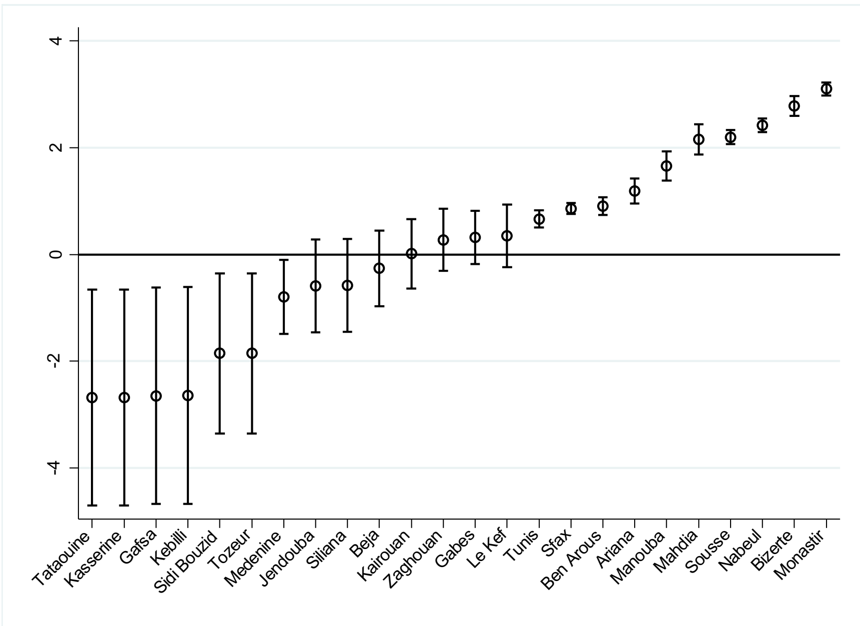



**Figure 2: Random effect estimates at governorate level (with 95% confidence intervals) for fully exporting firms**



# Appendix

## Table A.1: Descriptive statistics for firm-level and regional-level variables

| | Number of observations | Mean | Standard Error | Min-Max |
|---|---|---|---|---|
| ***Firm-level variables*** | | | | |
| Log of total factor Productivity (all sectors) using OP approach | 8199 | 6.965 | 0.751 | -0.525-10.805 |
| Log of total factor Productivity by sector (OP) | | | | |
| *Agro-food* | 948 | 7.149 | 0.923 | 1.307 - 10.166 |
| *Manufacture of non-metallic Products* | 611 | 6.812 | 0.896 | -0.525 - 9.888 |
| *Manufacture of machinery and equipment* | 202 | 7.035 | 0.908 | -0.254 - 9.277 |
| *Manufacture of transport Equipment* | 179 | 7.066 | 0.910 | 4.104 - 10.805 |
| *Metallurgy and metalworking* | 576 | 6.978 | 0.738 | 3.678 - 9.867 |
| *Electric, Electronics and electrical appliances industries* | 441 | 7.203 | 0.670 | 4.127 - 9.545 |
| *Chemical industry* | 458 | 7.381 | 0.799 | 4.075 - 8.958 |
| *Textile, clothing, leather and shoes industries* | 3910 | 6.860 | 0.622 | 1.269 - 10.722 |
| *Rubber and Plastics* | 377 | 7.031 | 0.709 | 2.601 - 9.136 |
| *Paper and Cardboard industries* | 370 | 6.959 | 0.785 | 1.487 - 9.484 |
| *Woodworking industry* | 127 | 6.749 | 0.772 | 4.698 - 9.087 |
| Log of total factor Productivity by export Status | | | | |
| *Non-exporting firms* | 3194 | 6.822 | 0.811 | -0.525 - 10.805 |
| *Partially exporting firms* | 1569 | 7.193 | 0.723 | 1.307 - 9.900 |
| *Fully exporting firms* | 3409 | 6.997 | 0.671 | 1.269 - 10.723 |
| Log of total factor Productivity by size groups | | | | |
| *Smaller than 50* | 353 | 6.800 | 0.836 | -0.525 - 10.805 |
| *[50, 200[* | 1704 | 7.006 | 0.666 | -0.254 - 10.723 |
| *Equal or more than 200* | 1187 | 7.263 | 0.614 | 1.269 - 10.107 |
| Export participation by size groups (partially or fully exporting firms) | | | | |
| *Smaller than 50* | 3184 | 0.366 | | |
| *[50, 200[* | 3673 | 0.742 | | |
| *Equal or more than 200* | 2278 | 0.907 | | |
| *All firms* | 9195 | 0.650 | | |
| Share of exports over total sales | | | | |
| *Smaller than 50* | 3160 | 0.237 | 0.400 | 0-1 |
| *[50, 200[* | 3665 | 0.582 | 0.467 | 0-1 |
| *Equal or more than 200* | 1281 | 0.690 | 0.435 | 0-1 |
| *All firms* | 8106 | 0.465 | 0.475 | 0-1 |
| Size of firm (employee) | 8199 | 125 | 223.7 | 6-4585 |
| Age of firm (year) | 7862 | 16.589 | 12.503 | 1-149 |
| ***Regional-level variables*** | | | | |
| Specialization index (LQ) | 1848 | 1.029 | 1.063 | 0 - 7.313 |
| Diversity index | 1848 | 3.778 | 1.205 | 1.468 - 6.301 |
| Relative diversity index | 1848 | 2.197 | 1.191 | 0.884 - 7.212 |

Authors' calculations based on the national annual survey report on firms (NASRF).

### Table A.2: Comparison and correlation between different TFP estimates obtained using different estimation approaches.

| | Descriptive statistics of TFP | | | | Correlation matrix | | |
|---|---|---|---|---|---|---|---|
| | Mean | Std | Min | Max | OLS | OP | ACF |
| Ordinary least squares (OLS) | 4.786 | 0.700 | -2.627 | 8.588 | 1 | | |
| Olley-Pakes (OP) | 6.965 | 0.751 | -0.525 | 10.805 | 0.932* | 1 | |
| Ackerberg-Caves-Frazer (ACF) | 4.524 | 0.701 | -2.869 | 8.332 | 0.998* | 0.912* | 1 |

### Table A.3: Elasticities of output with respect to labor and capital

| | Ordinary Least Square | Olley-Pakes | Ackerberg-Caves-Frazer |
|---|---|---|---|
| log of employment | 0.596*** | 0.567*** | 0.610*** |
| | (71.60) | (45.99) | (43.56) |
| log of capital | 0.432*** | 0.283 | 0.447*** |
| | (78.16) | (1.61) | (207.84) |
| Observations | 8199 | | |

Notes: t statistics in parentheses; *** p-value < 0.01.

### Figure A.1: Manufacturing firms' productivity distribution (in log) using alternative approaches