Firm productivity and FDI activity. Evidence from Danish industrialized firms

Boris Georgiev*

July 31, 2015

Abstract

The paper analyses the link between FDI and firm productivity. Using a longitudinal panel data for Danish industrialized firms, observed for a period of four years (2009-2012), it is found that multinational companies grow faster in terms of their productivity and also have higher levels of efficiency compared to non-multinationals. Productivity estimates are derived using a multilateral TFP Index along with semiparametric estimators by Olley and Pakes (1996) and Levinsohn and Petrin (2003). A major contribution is the productivity analysis between pairs of FDI host destinations. The results suggest that there are no significant differences in productivity growth among multinationals doing FDI in different host regions. However, significant differences in the levels of productivity between few pairs are found. The cross-regional comparison allows to construct a ranking of the FDI destinations by productivity. Only Danish firms with the highest productivity levels have subsidiaries in Africa and the Middle East, less productive ones have affiliates in Asia, South America, North America and Asia Pacific and the least productive MNEs predominantly have subsidiaries in Europe. The paper does not find support for the learning-by-doing hypothesis as the statistical test indicates that entrants are not growing faster than continuers on international markets. Finally, the FDI premium is estimated to be between 15.4% to 19% for multinationals depending on the specification.

1 Introduction

The topic of multinational activity has been gaining popularity since the 1960s when firms started gradually expanding their operations overseas. Nowadays, we are surrounded by products that are manufactured by big international corporations, which are not necessarily domestic companies. This trend towards globalization has continued and Denmark has been active in that process as well. Since the country is relying on exports and selling its products abroad, the activity of exporters and multinationals (MNEs) plays a very important role for the Danish economy. Thus the importance of international trade and establishing trade links is well known to Danish companies. Despite the importance of international markets for the Scandinavian country, companies that engage in Foreign

*School of Business and Social Sciences, Aarhus University. E-mail: bg@econ.au.dk. The author would like to thank Professor Frederic Warzynski and Professor Philipp J.H. Schröder for the useful comments and suggestions during the writing of this paper. Any mistakes are my responsibility.

Direct Investments are clearly a minority.[1] Taking a bird's eye view on the Danish economy, one would see that the majority of Danish firms only sell within Denmark, fewer export abroad and selected few engage in FDI.

The positive link between economic growth and FDI intensity has been well documented in the literature (Borensztein, De Gregorio and Lee, 1998). This is why investigating the Danish MNEs and their differences relative to other types of firms is of great interest. It is important because having more MNEs and being able to understand what drives their decisions to become international players, may have positive spillover effects to domestic firms Damgaard (2011). In the empirical literature the productivity-internationalization link has been well evaluated. On average multinationals and exporters are more productive than their purely domestic counterparts (Alfaro et al., 2004). This beneficial feature allows MNEs to afford the higher cost of serving a foreign market and establish subsidiaries abroad. This research paper follows the international trade literature in an attempt to understand what is the interplay between productivity and internationalization status by performing series of analyses, which outline a pattern that differentiates firms solely operating in Denmark from multinationals. FDI flows have been growing at a faster rate than other types of international transactions, which calls for more detailed understanding of what motivates firms to become multinational and understand how their decisions about international investments are shaped.

1.1 Danish outward FDI stock and flows

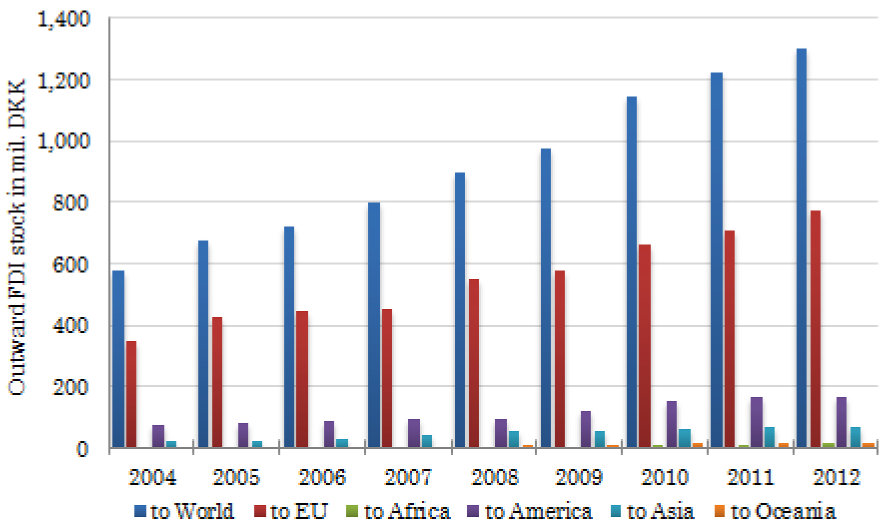
The FDI stock represents the cumulative value of Danish-owned assets, production facilities, firms and subsidiaries abroad. FDI flows are usually presented on a yearly basis and could be both negative or positive in value. For example, in a given year Danish firms may invest in new projects abroad or engage in mergers and acquisitions, which would show up as positive FDI flow for that year. On the other hand, a Danish firm could leave a given market and sell its subsidiaries in the foreign country, which would be seen as a divestment and have a negative sign in the FDI flow data.

Denmark, being a small open economy with active trade links, has seen its direct investments abroad grow faster than its export activity. Figure 1 shows the FDI stock accumulation across recipient regions. The positive development indicates that nowadays firms are actively looking to expand beyond their national borders. The total outward FDI stock in 2012 amounted to 1,304 billion DKK, which is an increase of more than 125% from 2004 when the investment stock was 579 billion DKK. On a regional basis the growth rates differ, which outlines the regions that have become more attractive among multinational Danish firms.

Overall, all areas have experienced a positive growth in the investment stock from 2004 until 2012. It is important to distinguish between growth and size of investments. The most notable increase in the FDI stock can be observed in Oceania (Asia Pacific), which includes Australia and New Zealand. The FDI stock by Danish firms has grown from 3 billion in 2004 to 18.1 billion DKK in 2012, corresponding to a staggering increase of 503%. This region is promising for Danish firms but on a worldwide basis it occupies a very small share of the Danish investment stock. The most important areas where firms have invested and expanded continuously are the EU, the Americas and Asia. The total stock growth in those regions from 2004 to 2012 is 122.7%, 127% and 200%, respectively. The European Union

[1] This activity is present when a firm decides to sell goods on foreign markets by affiliate production or subsidiary instead of exporting to that market.

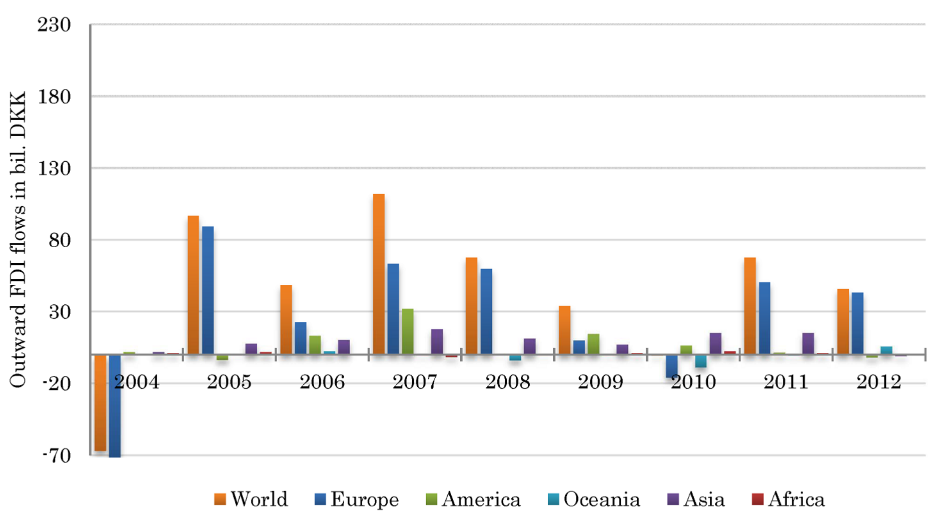
Figure 1: Outward FDI stock from Denmark by recipient region, 2004-2012



Source: Statistics Denmark, 2014; Variable: DNDIRA2

is very important for the Danish multinational firms as it represents close to 60% of the total FDI stock. This observation underlines the importance of understanding whether there are productivity differences across firms doing FDI in different continents. The table with the FDI stock data is available in Appendix C, Table 18. Figure 2 displays both positive and negative investments in the period between 2004 and 2012. After 2004 MNEs in Denmark have expanded their asset ownership by investing predominantly in Europe, followed by Asia and the Americas.

Figure 2: Outward FDI flows from Denmark by recipient region, 2004-2012



Source: OECD statistics, 2014; Variable: FDI flows by partner country

An interesting cyclical pattern can be seen in the investment flows. Since 2005 investments have been gradually declining on aggregate but in 2007 they rebounded again with a positive worldwide investment of more than 100 billion DKK by Danish firms. In the wake of the financial crisis outward FDI flows from Denmark have been declining steadily, reaching almost a balance between investments and divestments in 2010. In that year, some Danish firms partly withdrew their investments from European markets and those were redirected towards Asia and the Americas. Those two continents have seen positive yearly investments flows since 2004 until 2012. The cyclical pattern can be seen to continue with gradually declining FDI flows on aggregate in 2011 and 2012.

In light of the FDI development for Danish firms, the paper seeks to answer the following research questions:

1. Are there significant differences in the productivity levels between Danish multinationals and purely domestic companies?
2. Are productivity growth rates different among MNEs and purely domestic companies?
3. Do multinationals select different regions for FDI depending on their productivity levels?
4. Is there a productivity ranking of MNEs by host FDI region?

The answers to these questions offer several contributions to the literature. First, by using production function estimates, I find that there are significant differences in the productivity levels MNEs compared to purely domestic Danish industrialized firms. Second, the growth path of multinational firms is faster and differs significantly compared to purely domestic companies. Third, the carried out analysis shows that multinational firms select different regions for doing FDI and display significant productivity level differentials across geographical host areas. However, productivity growth rates among MNEs do not differ by FDI host region, which suggests that being a global player no matter where offers sizeable productiviy premia to firms relative to their purely domestic counterparts. Fourth, a natural ranking occurs based on this self-selection mechanism where ex-ante firms with the highest productivity levels can afford to setup affiliate operations in very distant regions. Finally, the productivity premia for being a multinational is estimated to be in the range of 15.4% - 19%.

The next section surveys the current state of the literature on multinational activity, trade and productivity. Section 3 introduces the methodology and Section 4 describes the data in detail. Section 5 offers the empirical analysis of productivity differentials between multinationals and purely domestic companies and presents an empirical framework for the FDI productivity premium. The last section concludes the paper and summarizes the main results of the study.[2]

2 Literature review

The most insightful theoretical contribution on the productivity ranking and self-selection into export markets is due to Melitz (2003) who explains in a theoretical framework why some firms export and others do not by pioneering the heterogeneous firms strand of the literature.[3] The key mechanism at place behind the theory is that firms are not all equally productive and some are more efficient in their operations than others in the economy. The prediction suggestes that as the overall productivity level increases, the survival productivity threshold increases which drives out the weakest (least productive)

[2]MNE and "multinational" are going to be used interchangeably hereafter.
[3]For an up-to-date review of the heterogeneous firms literature, see Melitz and Redding (2012).

firms out of the market in favor of the larger and more productive firms. The most efficient firms are able to cover the additional fixed costs of exporting so they self-select into serving these markets compared to domestic firms. In light of this theory the most productive firms self-select into the export markets or prefer to engage in FDI. MNEs that make foreign direct investments are the ones showing the highest productivity level since they are able to cover the fixed costs of establishing a production abroad and the less productive ones serve the foreign market from home.

In recent years there has been an effort to integrate theories of economic geography with "new" new trade theory. Ottaviano (2011) argues that firm heterogeneity has to be analyzed on a deeper level in conjunction with the choice of location of firms. This is important because space is heterogeneous in reality.[4] Thus, firms would try to make the natural choice to position themselves in areas which offer a comparative advantage. For example, a more beneficial location (advantaged) attracts bigger and more productive firms. This dilemma resembles heavily the popular chicken-egg question: Do more productive firms locate in more attractive locations or do they become more productive because they are in a more beneficial production location (spillovers among closely located firms)? This paper tackles this question by looking whether there is a self-selection into FDI in different locations or learning-by-investing effect at play. In this setup certain FDI regions are more attractive compared to others, which helps establish the differentiation across geographies.

Recent studies by Helpman, Melitz and Yeaple (2004); Greenaway and Kneller (2007); Bernard et al. (2003) shed light on the behavior of firms in the economy and their choice between exporting and doing FDI. The simplicity and true motivation behind engaging in FDI have received considerable attention. The main question has been whether firms invest abroad only because of market access considerations or they also consider vertical motives. A firm can now establish an export platform in a foreign country in order to serve the nearby locations. Alternatively it can choose to establish manufacturing facilities in order to optimize its value chain and import intermediate goods from its subsidiaries (Blonigen, 2005). Several studies have been exploring even more complicated patterns of FDI. For example, an interesting possibility is export platform FDI. Ekholm, Forslid and Markusen (2007) and Tintelnot (2013) investigate the cases where a multinational places FDI into a foreign country to serve as a production platform for exports to a group of closely located host countries. Furthermore, vertical interaction (or fragmentation) across the value chain is modeled by Baltagi, Egger and Pfaffermayr (2007) where affiliates of an MNE are shipping intermediate goods back to the home country in order to complete a finished good. In their model with many countries various characteristics such as country GDP, resource endowment and trade costs affect FDI choices along with the initial motivation whether to pursue horizontal, vertical or export-platform type.

The number of empirical studies investigating the link between exports, FDI and productivity has been on the rise in the past decade thanks to the new detailed datasets that follow firms across time on key performance variables such as sales, export share of the revenue, number of employees and FDI activity. These features allow researchers to dig much deeper into international behavior of companies. Such developments are not only important for empirical research and accumulation of knowledge but also for policy-making. In the following I refer to studies that report findings about the internationalization choices of firms and their productivity differences.

Brainard (1997) assesse the proximity-concentration trade-off that firms face when choosing to export versus doing FDI. Her results show that firms prefer to serve foreign markets via exports when

[4] Resources are unevenly distributed across countries.

there are clear cost advantages compared to the size of the foreign market. If the host market is not big enough then the high fixed costs of establishing foreign production abroad may not be justified. On the other hand, companies prefer to engage in FDI when the size and proximity to the foreign markets is more important. In that way a company can save on various variable costs associated with exporting such as transportation costs, levied duties, differences in the legal systems across states, country-specific fees and other impediments that may stem from language barriers. Empirical support can be found for the theoretical prediction that the share of exports increases if there are scale economies[5] but decreases in trade costs, distance to foreign country and size of the foreign market. Clearly there are the proximity-concentration forces that have opposite directions. Depending which one is prevalent would determine whether a foreign country is going to be served via exports or by establishing/acquiring production facilities abroad (FDI). The literature has acknowledged the importance of export fixed costs and the costs associated with FDI. Since only a minority of firms is able to cover these costs, the number of exporters and MNEs is a rather rare observation in trade data. This empirical pattern is reported for the U.S. (see Bernard and Jensen (1999)), for Mexico, Colombia and Morocco; (see Clerides, Lach and Tybout (1998)), for Canada (see Baldwin and Gu (2003)) and for Korea and Taiwan (see Aw, Chung and Roberts (2000)). These results are very important as they clearly underline the systematic relationship between firm characteristics and their export intensity and investments abroad. Therefore, MNEs and exporters are not randomly selected among the huge population of firms. They significantly differ in productivity, size, number of served markets, number of employees and investments.

One of the most often met theoretical predictions is that of productivity ordering of MNEs, exporters and non-exporters. The coexistence of exporters and non-exporters is addressed in Bernard, Jensen and Lawrence (1995) who note that there are significant differences in the productivity levels of the two type of firms. Girma, Kneller and Pisu (2005) use the concept of stochastic dominance to test the productivity ranking hypothesis. Their results are robust to various specifications of productivity and are consistent with the conjecture that trade patterns are affected by productivity heterogeneity among firms and industries. The empirically obtained ranking is as one would have expected from the Melitz model: the productivity distribution of multinationals dominates that of exporters and also that of non-exporting companies. Similar studies have been conducted in different countries, finding similar patterns in the data. For example, Delgado, Farinas and Ruano (2002) rely on the same methodological approach by using the Kolmogorov-Smirnov tests of stochastic dominance on the productivity distributions of Spanish exporters and non-exporting companies. The evidence suggest that the entire productivity distribution of exporters is to the right of the one for non-exporters.

The literature on MNEs and exporters has also benefited from several studies that investigate the self-selection versus learning-by-doing hypothesis. Finding a conclusive answer to this question is rather important as it may shed light whether the incentives for internationalization of firms, given by policy-makers, are effective. The key questions is whether the best firms (the most productive ones) play in a league of their own and self-select into foreign markets as they are able to compete internationally or firms just try their luck and learn while they try to export. In a study focusing on

[5] This very well known concept in economics is particularly important in the context of manufacturing firms. The reason is that companies can reduce the average cost per unit produced by having higher output and consequently distributing the fixed costs across more produced units. This is why the greater the economies of scale, the better it is to serve a foreign market from home as the goods demanded from abroad can be produced in the home factory, leading to lower average cost per unit.

self-selection of firms from Germany, United Kingdom and France, Temouri, Vogel and Wagner (2013) find evidence in support of the most productive firms choosing to enter export markets.

3 Methodology

In order to answer the posed research questions various methods are used to estimate as accurately as possible the productivity of Danish firms in the sample by applying several different estimators. Alternative measures of productivity are specified and discussed. The productivity differences among firms with different trade status are traced out using tests of stochastic dominance, applied to their productivity distributions.

3.1 Productivity measurement and estimation

In recent years firm-level data have become more available, which permits productivity analysis at very detailed firm level. The most popular approach in the literature has been to estimate a standard Cobb-Douglas production function of the following form:

$$Y_i = F(A_i, K_i, L_i) = A_i K_i^{\beta_k} L_i^{\beta_l} \tag{1}$$

which assumes that firms utilize two types of inputs: capital (K) and labor (L). Y_i is the firm's output measured either in terms of total revenue or value-added, A_i captures the efficiency with which a firm transforms the inputs into output, β_K and β_L represent output elasticity with respect to capital and labor. The efficiency term is of great interest because productivity cannot be observed by the econometrician in the data, but is probably known to the individual firm i. It can be thought of as a firm-specific productivity level. Equation 1 is theoretically motivated from the Solow growth model, which takes A as total factor productivity (TFP) in the economy.

Taking logs of Equation 1 yields the empirical specification of the production function on an industry level:

$$y_{it} = \ln(A_{it}) + \beta_k \underbrace{\ln(K_{it})}_{k_{it}} + \beta_l \underbrace{\ln(L_{it})}_{l_{it}} = \beta_k k_{it} + \beta_l l_{it} + \underbrace{\beta_0 + \psi_{it}}_{\ln(A_{it})} \quad \text{for } t = 1, ..., T \tag{2}$$

, where k_{it} is the logarithm of capital, K_{it}, l_{it} is the logarithm of labor input, L_{it} and the productivity term, $\ln(A_{it})$ can be decomposed into a time-constant and time-varying element, β_0 and ψ_{it}, respectively. The TFP measure is obtained as the residual from the estimation of a logarithmic Cobb-Douglas production function as in (2), which is the deviation between observed output and predicted output. Note that a time index is added in order to allow for temporal as well as cross-section treatment of the data. β_0 can be seen as a mean productivity in a given industrial sector and ψ_{it} are the shocks that make the individual productivity move away from the mean, which can be due to (i) measurement error in input and output accounting and (ii) unobserved factors that affect the output of a firm. The residual ψ_{it} can be decomposed further as :

$$y_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + \underbrace{\theta_{it} + v_i + \gamma_t + \eta_{it}}_{\psi_{it}} \tag{3}$$

, where θ_{it} is the productivity term[6] that is *correlated* with the input factors, v_i is the deviation from the mean industry efficiency of a given firm, which can be captured by using firm-level fixed effects as it is time-invariant and η_{it} is the error term that is *uncorrelated* with the inputs in the production function. Furthermore, a measurement error can be included in the residual, which is allowed to be correlated with the measurable inputs such as capital. In this specification one can allow for aggregate increases in average productivity across time, which may be due to improvements in technology and that affects all firms. Such effects can be captured by incorporating time dummies, γ_t into the regression model.

The productivity element, θ_{it} can be seen as the combined impact of firm's deviation from its own average TFP and also from the industry-wide average productivity in time period t only observed by the *firm* and not the *econometrician*. Marschak and Andrews (1944) state that the firm- and time-specific productivity shock, θ_{it} can be seen as the effort and will of a given entrepreneur that drives technological progress forward at a given point in time. The main concern with the above specification in 3 is that firms decide on their input factors based on the observed productivity shock, θ_{it} that the firm experiences. Thus, a company will decide on labor, l and capital, k depending on the observable TFP component. Estimating Equation 3 by OLS will result in biased estimates. Additionally, the rationale behind $Corr(k_{it}, \psi_{it}) \neq 0$ is that big firms have more complex structures and are more prone to errors in their measurement of "true capital" as they operate using more complex accounts in comparison to smaller firms (Eberhardt and Helmers, 2010).[7]

Consistent estimates of the parameters can be obtained by using semi-parametric techniques suggested by Olley and Pakes (1996) and Levinsohn and Petrin (2003), which have gained wide popularity in the TFP literature. These methods are very similar and manage to correct for the endogeneity bias by using different proxies for the firm's private knowledge of its productivity. Fixed-effect estimation can also be applied in order to obtain consistent estimates of the input parameters under a strict set of assumptions.

The methodology proposed by Olley and Pakes (1996) was the first to approach the simultaneity between output and variable inputs by using firm investments as a reliable proxy for the unobserved productivity shock. The estimated production function is highly similar to the one in Equation 3 but having material as additional input variable. The self-selection issue is targeted by estimating a probit model for the probability of survival. In their estimation strategy, Olley and Pakes (OP) use an extended production function which includes both variable and fixed inputs. The Olley and Pakes technique is hard to implement in reality as it requires positive investment flows in each period. Very often the applicability of the method is data-driven as very few firm-level datasets report investments due to their lumpiness. Changing investments involves high adjustments costs, which makes the use of the variable as a proxy rather difficult. Section 4 is mentions that the ORBIS database contains very few investment entries, which required the use of the perpetual inventory method in order to obtain investment levels from the capital dynamics. The obtained TFPs of companies are not comparable across industries due to technological differences in their production processes.

Alternatively, Levinsohn and Petrin (LP) provide a solution to the transmission problem by using

[6] The same notation standard is kept for the "productivity" term, θ as in the theoretical model, introduced in section **??**.

[7] Consider that the true capital stock, denoted by $\dot{k}_{it}$ is measured with an error ξ_{it}, so that $k_{it} = \dot{k}_{it} + \xi_{it}$. Then, the simple production function $y_{it} = \beta_0 + \beta_l l_{it} + \beta_k \dot{k}_{it} + \eta_{it}$ is estimated as: $y_{it} = \beta_0 + \beta_l l_{it} + \beta_k k_{it} + (\eta_{it} - \beta_k \xi_{it})$. This reveals the negative correlation between capital and the residual in the brackets.

intermediate inputs as a proxy. To apply the technique, data on materials used in the production process are required to account for the internal firm knowledge of its productivity. This method is more desirable than OP because intermediate inputs are less costly to adjust. Additionally, data on materials are more widely available in firm-level datasets. This relaxes the heavy data requirements imposed by the OP methodology. The benefit of the LP method is that firms almost always have positive values for materials and other intermediate inputs, whereas investments are not only always positive in a given time period. Thus, the LP approach avoids the truncation issue that could arise from using only positive investments levels. A further argument for the LP approach is that the value of materials is always positive, since if it is zero, then firms simply cannot operate on the market.

The empirical setting departs from the Cobb-Douglas production function as the one in Equation 2, where the inputs are log-transformed. The demand for additional variable - intermediate inputs, m_{it} is assumed to be dependent on the state variables capital, k_{it} and firm productivity, θ_{it}, such that: $m_{it} = m_{it}(k_{it}, \theta_{it})$. Levinsohn and Petrin (2003) shows that the intermediate input demand, m_{it} is monotonically increasing in the efficiency term, θ_{it}. This implies that firms observing higher productivity levels inevitably use more intermediate materials, which may not always be plausible in reality. Similar to the OP inversion, the unobserved TFP component can be represented as a function of capital and intermediate inputs:

$$\theta_{it} = m_{it}^{-1}(k_{it}, m_{it})$$

, where now the productivity term is a function of two observed variables. It is still assumed that the productivity follows a first-order Markov process. When using the value-added as a dependent variable, the equations of interest take the following form:

$$va_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + \beta_m m_{it} + \underbrace{\theta_{it} + \eta_{it}}_{\ln(A_{it})} = \beta_l l_{it} + \gamma_{it}(k_{it}, m_{it}) + \eta_{it} \tag{4}$$

, where $\gamma_{it}(k_{it}, m_{it}) = \beta_0 + \beta_k k_{it} + \beta_m m_{it} + \theta_{it}(k_{it}, m_{it})$ is approximated by a third-order polynomial. Inserting the polynomial in Equation 4 allows to obtain consistent estimates for β_l and $\gamma_{it}(\cdot)$ up to the intercept. The second estimation stage identifies the capital coefficient, β_k by using the estimated values for the variable $\hat{\gamma_{it}}$ and minimizing the sample residual of the production function.[8]

Under the strong assumptions the fixed effects estimator can be implemented as well in the production function estimation. The unobserved firm characteristics, v_i should be constant across time in Equation 3 so that they can get eliminated by the transformation.[9] An important precondition for the application of the FE estimator is the time-variation of the variables of interest. If they are constant, the parameters cannot be identified as they will be wiped away due to the FE transformation. In the current setting, it can be assumed that firms change their inputs across time with labour being the more flexible input compared to capital. For this method to bring unbiased results, an assumption that there are no other productivity shocks that affect the individual firm are required. This implies that $\theta_{it} = 0$. This setting means that unobserved productivity differences among firms in a given

[8]The exact derivations of the LP method can be found in Levinsohn and Petrin (2003) and Petrin, Poi and Levinsohn (2004).

[9]Since the average of v_i across a given panel is equal to: $\bar{v_i} = \frac{\sum_t^T v_i}{T}$, then the demeaned result of this unobserved heterogeneity is: $\tilde{v_i} = v_i - \bar{v_i} = 0$. As a result the transformed production function is now without this element, which allows for the consistent parameter estimation. For more information on fixed effects estimation, see Wooldridge (2010).

sector are constant over time. Under strict exogeneity assumption where the explanatory variables are uncorrelated with the error term across all periods: $Corr(x_{jit}, u_{it}) = 0$, the fixed effects (within) estimator is unbiased and consistent Wooldridge (2010). Appendix E offers a summary of the used estimators by outlining their benefits and drawbacks.

The Multilateral TFP index approach is implemented for the entire productivity analysis of multinationals and purely domestic firms. It is computed by combining output and the cost shares of inputs. The index approach is able to capture differences across companies in technology, size, age and industry. The main interest of applying this methodology is to compare TFP differences between firms in a given year and also track the shift in the productivity distribution across time. The multilateral index specifies a hypothetical reference company, which has a turnover equal to the average revenues of all firm observations in the sector, the mean amount of capital, labor and cost-shares. The efficiency of an individual firm is evaluated vis-à-vis the mean efficiency in its industry in a given base time period. The TFP of the reference hypothetical company in a given industry is computed by NACE 2-digit sector. The original Index TFP measure takes the following form:

$$\begin{aligned} \ln(TFP_{it}) \quad = \quad & \left(\ln(Y_{it}) - \overline{\ln(Y_t)}\right) + \sum_{s=2}^{t}\left(\overline{\ln(Y_s)} - \overline{\ln(Y_{s-1})}\right) \\ & - \sum_{m=1}^{n} \frac{1}{2}\left(s_{mit} + \overline{s_{mt}}\right)\left(\ln(X_{mit}) - \overline{\ln(X_{mt})}\right) \\ & - \sum_{s=2}^{t}\sum_{m=1}^{n} \frac{1}{2}\left(\overline{s_{ms}} + \overline{s_{m,s-1}}\right)\left(\overline{\ln(X_{ms})} - \overline{\ln(X_{m,s-1})}\right) \end{aligned} \tag{5}$$

, where $\ln(Y_{it})$ is the log-transformed output of company i at time t, which in this specification is value-added, $\overline{\ln(Y_t)}$ is the log of the average output in the industry for the given year t, s_{mit} is the cost share of input m of firm i at time t, $\overline{s_{mt}}$ is the average cost share of input m in the industry during time period t, $\ln(X_{mit})$ is the log of the amount of input X_m and $\overline{\ln(X_{mt})}$ is the average level of input X_m in the industry for $m = 1...n$, where m is the number of inputs used in the production. The proposed computation is used as a proxy measure for θ in the theoretical models. Heterogeneity in the production processes across firms are also captured by the input cost shares via the s_{mit} factors. The multilateral TFP methodology is developed in the work of Caves, Christensen and Diewert (1982), Aw, Chen and Roberts (2001) and Caves, Christensen and Tretheway (1981), where the latter authors apply the approach for first time to the airline industry in the USA. Similar to the regressions, the productivity of the firm is measured as the difference between the output and the inputs that enter the production process. This proportion of output, unexplained by the inputs can be referred to as efficiency. By adopting the Index TFP measurement, productivity across firms can be compared.

3.2 Kolmogorov-Smirnov tests of stochastic dominance

The method is adopted with respect to the productivity distributions of firms depending on their international trade status The test is more desirable than a simple mean comparisons between firms as it also tests the stochastic dominance of the entire distribution and not only the first moment. The Kolmogorov-Smirnov (K-S) approach is popular in the empirical trade literature as it allows to test the productivity ranking derived from theoretical models. Girma, Görg and Strobl (2004) apply the K-S methodology on Irish data and Delgado, Farinas and Ruano (2002) on Spanish firm data in an empirical analysis of productivity ranking of firms. The results in both studies support the efficiency ordering suggested by the theoretical model of Melitz, Helpman and Yeaple (2004).

This method is applied on the productivity levels of purely domestic firms vs. MNEs. This method is used to shed light on the question whether firms doing FDI in a particular region of the world have higher average productivity than firms investing in a reference region. In order to establish a first-order stochastic dominance requires that efficiency distribution of the more productive firms lies to the right of the less productive ones. Finding a difference in the means of two distributions does not imply that there is stochastic dominance, whereas the reverse is true (Conover, 1999).

Let $\ln(TFP)$ be defines as the logarithm of the total factor productivity and $G(\ln(TFP))$ and $C(\ln(TFP))$ two cumulative distributions, respectively. The former one is the distribution of purely domestic producers and the latter one is for firms that are involved in FDI. Depending on data availability, one could test the equality of various distributions of productivity such as: non-exporters vs. exporters; MNEs vs. non-exporters and MNEs vs. exporters. The two-sided test can be specified as:

$$H_0 : C(r) - G(r) = 0 \ \forall r \in \mathcal{R}$$
$$H_1 : C(r) - G(r) \neq 0 \text{ for some value of } r \in \mathcal{R}$$

, where under the null hypothesis there are no differences between the two distributions. If the K-S test statistic[10] cannot reject H_0 under plausible levels of statistical significance, then stochastic dominance of $C(r)$ over $G(r)$ cannot be established. The alternative hypothesis states that one of the cumulative distributions stochastically dominates the other. The one-sided test has a modified H_1 where the hypothesis is that $C(r)$ strictly dominates $G(r)$, meaning that: $C(r) - G(r) > 0$. With the one-sided test after rejecting the null hypothesis, stochastic dominance of $C(r) > G(r)$ can be asserted.

4 Data

The dataset has been constructed by extracting disaggregated firm-level data from the international database ORBIS, assembled by the Bureau van Dijk (BvD).[11] Several other studies, such as Faggio, Salvanes and Van Reenen (2010) and Fons-Rosen et al. (2012), rely on data from BvD for productivity analysis due to its accessibility. The used dataset contains only firms registered in Denmark. The sample includes only economically active and industrialized companies in period the 2009-2012. Firms are considered active if they had at least one employee in 2011 and 2012. Having access to a longer panel is clearly desirable as it would allow to capture the productivity evolution of firms before the crisis. In the current situation only the post crisis behavior can be discussed and tracked. Observations that have missing entries or negative values for revenues and cost of goods sold have been removed from the dataset.

[10] The one-sided test statistic is: $K = \sqrt{(n \cdot m)/N} \cdot \max_{1 \leq i \leq N} |C_n(r_i) - G_m(r_i)|$, where n is the sample size of distribution $C(r)$, m is the sample size of $G(r)$ and N is the sum of both sample sizes n and m.

[11] The database can be accessed here: https://orbis.bvdinfo.com/. Firms in Denmark report to the Commercial Register. KOB A/S is the Danish data provider, which collects the information on a local level and sells it to ORBIS.

From a data coverage point of view, ORBIS is far from ideal. The reason is the missing entries on some key variables such as investments, material costs, R&D, revenues, cost of goods sold and export revenue. A very small fraction of the firms have information on these crucial variables. Furthermore, data limitations do not allow to distinguish purely domestic firms from exporters as export's share in the revenues is not reported. The possible analysis is to test the productivity differences between a group of firms that do not have subsidiaries abroad (do not engage in FDI) and those that invest in foreign facilities (FDI doers). Despite the stressed caveats, the data from BvD remains widely accessible for public research and some of the mentioned issues can be circumvented.

For some of the missing variables, approximate measures are built, which are explicitly discussed in Subsection C.1. The final dataset contains information on firm size, industry (NACE Rev.2 classification),[12] company name, date of establishment, age in days and years, subsidiaries and a wide range of financial and accounting variables. An important delimitation of the project is that exporters are excluded from the analysis, because they cannot be identified in the data. The source of the dataset does not distinguish between exporters and non-exporters for Denmark. Since several productivity estimates are obtained, the Index efficiency has been used as a core measure in the analysis. The rest are reported for robustness purposes, but the entire analysis is not repeated for each productivity measure due to space constraints.

4.1 Descriptive statistics

The summary statistics of the most important variables, needed for the productivity analysis are presented in Table 3. Since there are numerous industries, some of them are very weakly represented in the datasets. Industries such as public administration and defence, mining and real estate are clear examples of such sectors.

Companies from 19 industries are included in the sample but they are not uniformly distributed as can be seen in Table 1. Only firms with consistent information across the four years are kept in the reported table. Several industries are dominant in the dataset, such as "Manufacturing" and "Wholesale and retail trade", which is to be expected. Table 1 also allows to see within each industry what is the size of the firms. 83.29% of the manufacturing companies are large or very large. The number for Wholesale and retail is 61.55%. On the other hand, sectors such as food services, social and public works are predominantly represented by small companies and occupy a very small fraction of the dataset. It is important to stress that the sample of firms is not fully representative of the population of all industrialized Danish firms, due to reporting quality of smaller firms, which is discussed in subsection C.2 in the Appendix. Data from Statistics Denmark, suggests that the number of firms present in balanced dataset represent less than 0.8% of the total company population in the country. It is evident that there is a data selection towards large and very large companies. This is why the productivity estimates shall be interpreted with attention about their external validity.

The attrition problem among smaller firms is evident as after cleaning the data their share declined in the final dataset. Table 2 illustrates that the balanced panel contains 21.47% of firms being very large, 39.36% are large, 20.58% are medium enterprises and finally only 18.59% are small. The pictures is different when looking at the unbalanced panel. Small firms represent more than 40% before 2012

[12]The NACE industrial classification is widely used to arrange various sectors of the economy. For more information about the format, see the online Eurostat publication "Business economy by sector".

Table 1: Firm distribution by size and industry (2009-2012) - Balanced panel

Industries:	Small	Medium-sized	Large	Very large	Observations:	% of total:
1 Agriculture, forestry and fishing	35.29%	41.18%	11.76%	11.76%	68	0.8%
2 Mining and quarrying	0%	14.29%	42.86%	42.86%	56	0.7%
3 Manufacturing	4.47%	12.24%	49.41%	33.88%	1,700	20.1%
4 Electricity, gas, steam and air conditioning	4.35%	17.39%	47.83%	30.43%	92	1.1%
5 Water supply, sewerage, waste management	11.76%	35.29%	29.41%	23.53%	68	0.8%
6 Construction	37.21%	26.74%	29.65%	6.40%	688	8.1%
7 Wholesale and retail trade, repairs	15.35%	23.10%	41.99%	19.56%	2,372	28.0%
8 Transportation and storage	12.77%	19.15%	38.30%	29.79%	376	4.4%
9 Accommodation and food service activities	24.24%	30.30%	33.33%	12.12%	132	1.6%
10 Information and communication	23.16%	29.94%	33.33%	13.56%	708	8.4%
11 Financial and insurance activities	3.52%	3.52%	52.82%	40.14%	568	6.7%
12 Real estate activities	0%	0%	0%	100%	4	0.0%
13 Professional, scientific and technical activities	38.31%	22.98%	26.61%	12.10%	992	11.7%
14 Administrative and support service	15.48%	19.05%	42.86%	22.62%	336	4.0%
15 Public administration and defence	0%	0%	0%	100%	4	0.0%
16 Education	20%	40%	40%	0%	20	0.2%
17 Human health and social work activities	64.10%	25.64%	10.26%	0%	156	1.8%
18 Arts, entertainment and recreation	16.67%	50%	27.78%	5.56%	72	0.8%
19 Other service activities	50%	12.50%	37.50%	0%	64	0.8%
Total observations:					8,476	

Sources: ORBIS and author's own calculations based on the constructed dataset
Note: The distribution in the table is based on the final dataset, which is a balanced panel only containing firms without missing data on revenues, cost of goods sold, number of employees, total tangible assets, capital and industry classification.

where their share drastically drops due to missing values on some key explanatory variable for that year.

The descriptive analysis is supplemented with information on multinationals and their purely domestic counterparts. A first observtation is that non-MNEs are prevalent in the dataset. Similar empirical studies find that firms in international trade are the exception rather than the rule. In the balanced database, out of the 8,476 observations, 5,965 are purely domestic firms and 2,511 are of firms that are involved in FDI abroad. In percentage terms, in 2009 28.45 % of the companies were MNEs and 71.55% did not have any foreign affiliates. In 2010, MNEs accounted for 29.78% and non-MNEs represented 70.22% of the sample. The year 2011, saw an increase in firms engaging in FDI as 30% of the companies had at least one affiliate. In the last year of observation, purely domestic companies are 69.8% of the sample and the MNEs are 30.2%. This development suggests that some firms converted from purely domestic into multinational companies, since the share of companies engaging in FDI has been increasing within those four years.

Next, the key variables for the productivity estimation are discussed in light of the different international status of firms. Table 3 shows that on average purely domestic firms are smaller than multinationals, have lower revenues, possess less capital (tangible assets) and generate lower value-added. These preliminary observations are inline with the literature, where Manova and Zhang (2012) and Eaton, Kortum and Kramarz (2004) report similar observations for a panel of Chinese and French firms, respectively. Similar static observations between MNEs and domestic firms are also found in De Loecker (2007), who uses Slovenian firm level data and in Arnold and Hussinger (2005) who rely

Table 2: Firm segmentation by size and year after data cleaning (2009-2012)

	Balanced panel				Unbalanced panel			
Size \ Year	2009	2010	2011	2012	2009	2010	2011	2012
Small	18.59%	18.59%	18.59%	18.59%	42.82%	41.58%	41.29%	24.49%
Medium-sized	20.58%	20.58%	20.58%	20.58%	20.19%	20.71%	20.51%	22.20%
Large	39.36%	39.36%	39.36%	39.36%	24.16%	24.51%	25.00%	35.23%
Very large	21.47%	21.47%	21.47%	21.47%	12.82%	13.21%	13.20%	18.08%
Total:	100%	100%	100%	100%	100%	100%	100%	100%
Observations:	2,119	2,119	2,119	2,119	4,110	4,709	5,008	3,540

Note: The statistics are based on the cleaned dataset, corresponding to stage 3.

on German data for manufacturing companies.

The average revenue for both domestic and MNEs has been increasing across the four years. This reflects the partial recovery from the financial crisis and positive growth in sales. In the last reported year - 2012, the average turnover of a purely domestic company was close to 277 million DKK with big variation across firms as indicated by the standard deviation. This dispersion is even more pronounced for multinational companies where the average revenue was close to 3.4 billion DKK for 2012. This is to be expected because firms involved in trade have access to more markets and thus can sell in greater quantities.

The value-added for firms operating on international markets is greater by a factor of 10 than those serving only the local Danish market. Without surprise, MNEs employ much more people as indicated in Table 3. Except a small dip in employment in 2009, overall firms have been hiring more people as the average has been increasing gradually for both types of firms. The average domestic company employs close to 100 people, which is approximately 20 times less than a multinational firm. These stylized observations shed light on the big importance of understanding firms that are actively involved in the international markets. Obviously, they are crucially important for the economy because they generate more employment, revenues and value-added. By serving diverse markets around the world a firm could diversify its revenue streams, which reduces the potential impact of local negative shocks.

Table 4 shows the number of firms that had at least one subsidiary in Europe, the Middle East, Asia (excl. the Pacific region), Asia Pacific, North America, South American and Africa by year. The presentation is split into two parts. The column "Firms with subs. in *only one* region" identifies the number of firms that *only* have subsidiaries in the given region and nowhere else. In 2012, 324 companies in Denmark had affiliates in Europe and no other affiliates in other regions. This number has stayed fairly constant in Europe despite the small increase in 2010. The region Asia Pacific refers to countries such as Australia, New Zealand and the surrounding small islands around the area. More firms now have affiliate presence in the Asia Pacific region. Across the Atlantic ocean, there not many Danish companies that do FDI only in North or South America. The number of firms doing FDI in those regions has declined in recent years. Similar observation can be made for Africa, where the number of firms doing FDI has stayed very stable.

The second column, named "Subs. in several regions" features the number of companies that have at least one subsidiary in the given geographic region but *also* potentially in other regions. In

Table 3: Full descriptive statistics for production function variables of purely domestic firms and MNEs (2009-2012)

	Variable		Revenues		Value-added		N. of employees		Tangible assets	
	Year	Obs.	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.
Domestic firms	2009	1,516	245,699	851,876	65,763	544,054	96.42	234.22	57,806	502,551
	2010	1,488	250,073	885,700	72,846	626,755	94.67	247.91	59,431	461,888
	2011	1,482	276,277	1,110,389	80,363	813,950	97.02	266.00	61,520	370,385
	2012	1,479	276,589	917,489	73,033	507,935	97.80	273.45	56,574	287,761
Total firm observations: 5,965										
Multinationals	2009	603	2,746,688	12,900,000	922,128	5,291,986	2,047	20,459	1,139,866	10,400,000
	2010	631	2,921,311	14,700,000	1,090,471	6,911,989	1,976	20,861	1,143,834	10,600,000
	2011	637	3,148,793	15,400,000	1,135,593	6,879,228	2,057	21,881	1,174,352	11,100,000
	2012	640	3,410,373	16,600,000	1,178,816	6,974,518	2,081	21,981	1,184,002	11,300,000
Total firm observations: 2,511										

Note: The monetary values are in '000s of DKK.

that sense, the first column "Firms with subs. in *only one* region" is a subset of "Subs. in several regions". In 2012 there were 573 multinational Danish firms that had affiliates in Europe. Out of those, 249 (calculated as $573 - 324$) companies also have affiliates in other regions of the world.[13] Looking across all the geographical regions, it is evident that most multinationals do not operate only in one region. Instead, once a firm enters the international market, it sells its products to more than one continent, meaning that they show up in several rows of Table 4. This is why in the column "Firms with subsidiaries in: *A given* region" the sum of all firms exceed the number of uniquely identified companies in the dataset, which are 2,119.

Europe appears as the most preferred location for international expansion of Danish firms because of the market proximity, harmonized laws and ease of access. Referring to the last year of observation, North America is also highly preferred choice for FDI flows. Even though in 2012 there were very few companies doing FDI *only* in North America, there are 175 (calculated as $191 - 16$) MNEs that have subsidiaries in North America and elsewhere around the world.

Asian markets have attracted the interested of Danish multinationals. Combined together, Asia and Asia Pacific, in 2012 had 290 Danish multinational companies operating there. Out of these, 24 companies did FDI solely in the Asian region, whereas the remaining 266 firms have subsidiaries in other areas as well. Especially in continental Asia, the number of firms with affiliates has more than doubled from 42 in 2009 to 104 in 2012, which underlines the attractiveness of emerging economies as manufacturing hubs. This region displays also the highest growth in terms of firms that established subsidiaries in that region.

Table 5 shows the firms that engage in FDI by their size. It is clear that small and medium companies are rarely establishing foreign affiliates. Those that do, prefer closer regions to Denmark such as Europe. For example, 6.4% of the medium-sized companies in our database had at least one

[13] To obtain the number of firms operating in the given row geographic location and elsewhere around the world, one has to calculate: "Firms with subs. in a *given* region" - "Firms with subs. in *only one* region" by year.

Table 4: Location and number of firms with subsidiaries by geographic region (2009-2012)

		Number of DK firms with subsidiaries in:							
		Only one region				*A given* region			
Region	Year	2009	2010	2011	2012	2009	2010	2011	2012
Europe	Number	324	331	325	324	549	571	576	573
	% of all firms	15.3%	15.6%	15.3%	15.3%	25.9%	26.9%	27.2%	27.0%
Middle East	Number	1	3	2	2	43	47	56	52
	% of all firms	0.0%	0.1%	0.1%	0.1%	2.0%	2.2%	2.6%	2.5%
Asia excl. Pacific	Number	2	7	8	6	42	85	101	104
	% of all firms	0.1%	0.3%	0.4%	0.3%	2.0%	4.0%	4.8%	4.9%
Asia Pacific	Number	13	14	13	18	149	173	176	186
	% of all firms	0.6%	0.7%	0.6%	0.8%	7.0%	8.2%	8.3%	8.8%
North America	Number	19	16	14	16	187	189	191	191
	% of all firms	0.9%	0.8%	0.7%	0.8%	8.8%	8.9%	9.0%	9.0%
South America	Number	2	3	2	1	60	64	71	73
	% of all firms	0.1%	0.1%	0.1%	0.0%	2.8%	3.0%	3.4%	3.4%
Africa	Number	4	4	5	4	41	40	45	49
	% of all firms	0.2%	0.2%	0.2%	0.2%	1.9%	1.9%	2.1%	2.3%
Purely domestic	Number	1,517	1,489	1,483	1,480	1,517	1,489	1,483	1,480
	% of all firms	71.6%	70.3%	70.0%	69.8%	71.6%	70.3%	70.0%	69.8%
Total number of firms		2,119	2,119	2,119	2,119	2,119	2,119	2,119	2,119

Note: The table is based on the balanced and cleaned panel of Danish companies

subsidiary in Europe. For geographic areas farther away, very few firms undertake FDI among the small and medium companies. The picture changes drastically for large and very large establishments. For example 250 Danish companies that are classified as large had affiliate operations in Europe, which is close to 30% of the 834 large firms in the database. The number is even higher for very large companies, where 63% of the biggest corporate entities in Denmark had subsidiaries in the European area; 9.1% in the Middle East, 14.1% in continental Asia, 24.7% in Asia Pacific, 20% in North American, 11.5% in South America and finally only 7.6% in Africa.

The data indicate that there is an outlined tendency for large and very large companies to do much more FDI and be present in more markets than smaller and medium-sized companies. Furthermore, large and very large firms represent 60.9% of the entire sample, which signals potential data selection issues.

Finally, some descriptive statistics are offered for the number of subsidiaries by region and by year that Danish companies have. Figure 3 focuses on the average number of subsidiaries that those firms have around the world. The previous table suggested that Europe is one of the most preferred destinations with more than 500 Danish companies having foreign affiliate operations. An interesting question to answer is among these firms that are serving international markets, what is the average number of subsidiaries that each firm has. As expected, the average Danish MNE with operations in Europe has 5.8 subsidiaries on average. This number is higher than the after-crisis value but has declined since 2010 when the peak was at 6.5 establishments abroad per company as indicated by

Table 5: Internationalization status of firms by size and FDI host region (2009-2012)

		Company size			
Region	Year	Small	Medium	Large	Very large
Europe	Number of firms	3.2	27.7	249.5	286.7
	% of all firms	0.8%	6.4%	29.9%	63.0%
Middle East	Number of firms	0.0	0.3	0.8	41.3
	% of all firms	0.0%	0.1%	0.1%	9.1%
Asia excl. Pacific	Number of firms	0.0	0.3	18.5	64.2
	% of all firms	0.0%	0.1%	2.2%	14.1%
Asia Pacific	Number of firms	0.0	3.7	54.7	112.5
	% of all firms	0.0%	0.9%	6.6%	24.7%
North America	Number of firms	0.0	4.5	66.7	91.0
	% of all firms	0.0%	1.0%	8.0%	20.0%
South America	Number of firms	0.0	1.3	13.3	52.5
	% of all firms	0.0%	0.3%	1.6%	11.5%
Africa	Number of firms	1.0	1.7	6.3	34.8
	% of all firms	0.3%	0.4%	0.8%	7.6%
Total number of firms by category:		394	436	834	455
Percent of sample:		18.6%	20.6%	39.4%	21.5%

Note: The table is based on the balanced and cleaned panel of Danish companies

sub-figure 3a. Table 4 suggested that there has been a partial retaliation from the European market. An important reason for this development has been the slow economic recovery in Europe, which has not justified a further expansion of firms in the region.

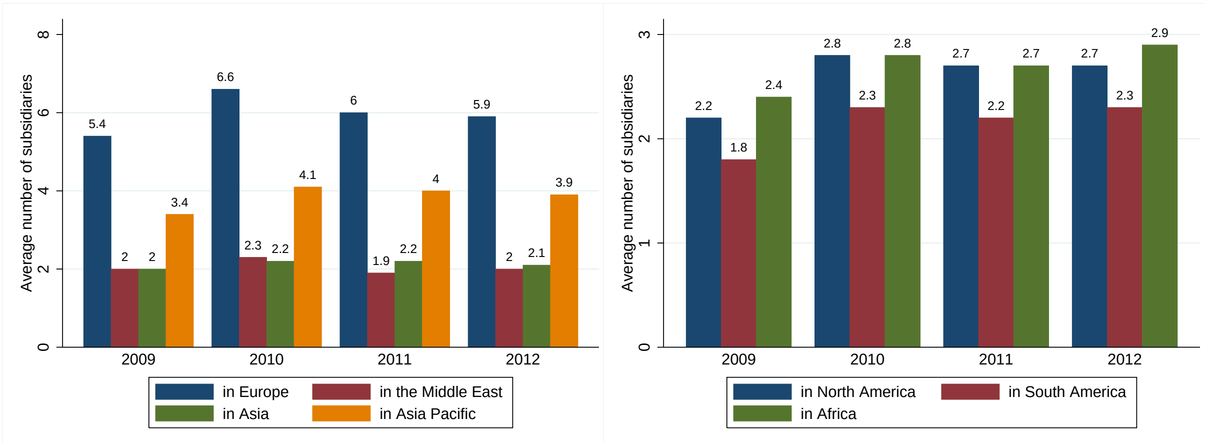
The number of subsidiaries per firm in the Middle East and Asia have stayed close to two and not much variation is seen across the four years. At this granular level of treatment there are differences between Asia Pacific and Asia (also referred to as Continental Asia). The number of affiliates per firm in the Asia Pacific region is almost twice as the average number of subsidiaries in Asia. The peak in expansion has been around 2010, where the average Danish firm undertaking FDI in the Asia Pacific area had around four subsidiaries.

Sub-figure 3b reveals the development in average number of subsidiaries per firm for the remaining three geographical regions. In 2009 Danish companies had on average 2.2 subsidiaries each in North America, 1.8 in South America and 2.4 in Africa. After the wake of the financial crisis these numbers have increased for all the three regions. In 2012, MNEs with presence in North and South America and Africa had on average 2.7, 2.3 and 2.8 subsidiaries, respectively.

In conclusion to the data section it is worth mentioning several aspects of the descriptive statistics. First, the final sample of 2,119 companies comprises of 19 industries, where “Manufacturing”, “Wholesale and trade retail” and “Professional and technical activities” all together account for more than 50% of the observations. Second, large and very large Danish firms tend to be engaged in foreign direct investments, whereas smaller companies are mostly operating only within Denmark. The next chapter puts the data to the test by carrying out multiple tests and statistical analysis in order to give conclusive answers to the research questions.

Figure 3: Average number of subsidiaries by year and geographic region

(a) Europe, the Middle East, Asia and Asia Pacific (b) North and South America and Africa



5 Empirical analysis

The first subsection presents the productivity estimates of different types of firms and constructs a ranking. The following section builds an econometric model that tests whether firms engaged in FDI enjoy a productivity premium compared to firms that stay purely domestic. First, using the TFP Index and semi-parametric estimators, the hypothesis that firms operating abroad dominate domestic companies in terms of productivity is confirmed. Second, the ranking of the different FDI recipient regions around the world are assessed. Third, the hypotheses that MNEs grow faster and have higher productivity levels are tested by specifying an econometric model. For a simple illustration of the theoretical guidance behind the testable hypotheses, Appendix A offers an exposition of the differences in the profit function among MNEs, exporters and purely domestic firms.

5.1 Productivity analysis

First, several measures of TFP[14] are presented and the results of these estimations are discussed. Four different TFP measures are computed in order to assess the robustness of the results. The residuals from an OLS regression of a production function are obtained for comparison purposes. The correlations of the four efficiency constructs are presented in Table 25 in Appendix D. Since TFP can be defined in several ways, the following measures have been included: Labor Total Factor Productivity (L_TFP), Index-based TFP (I_TFP), TFP by Levinsohn and Petrin (LP_TFP), TFP obtained by regular OLS regression (OLS_TFP) and finally TFP estimates obtained from Olley and Pakes (OP_TFP).

The Index TFP is highly desirable as it takes into account the different cost share of inputs by industries, which in a way takes into account firm heterogeneity. In the other efficiency approaches this is not taken into account. One possibility is to estimate the Levinsohn and Petrin (2003) and Olley and Pakes (1996) by industrial sector. In that way the obtained coefficients will be specific to the given industry but the downside is the little number of observations, which may hinder the

[14]The term TFP is used interchangeably with productivity and efficiency.

significance of the estimates. It has been decided to obtain LP estimates from regressions ran on an industry basis. Then the results are appended together for each estimation method. The efficiency terms can be obtained for all observations when using the Index TFP method, the labor TFP and LP (since intermediate inputs are always positive and available). For more than half of the sample, the logarithm of investment values is missing due to non-positive investment levels, which complicates the application of the OP productivity estimation approach. The key criteria for productivity evaluation is the Index-based TFP (I_TFP), which makes use of the total number of observations, does not pose specific econometric requirements and takes into account industry specific factors such as input cost shares.

The estimates from the LP, OP and OLS estimations are compared and discussed. The OLS and OP results are ran on the total sample due to data constraints. The LP and FE estimations are performed on sectoral level, which allows to capture inter-industry heterogeneity. The OP estimates are performed on the unbalanced panel, simply because the logarithm of investments is missing for more than 78% of the observations. In the process of obtaining values for investments, one year of observations is lost when applying the OP method. A further reduction in the observations arises because of the many firms having negative investments (less capital than the previous year), which means that the logarithm is undefined for these negative values. After the crisis a lot of companies had a reduction in their tangible assets, which suggests that investments have been negative from period to period for a big number of firms. The sample on which investments can be used as a proxy is drastically reduced. That is the primary reason why OP estimates could not be obtained on an industry level, but instead the production function estimator is applied on the entire sample. Around 500 observations have been dropped from the cleaned balanced dataset as for some firms the labor and capital input cost shares exceeded 1, which is not possible. These firms were removed and the panel was rebalanced again. The obtained sample that is used for the productivity analysis and K-S tests contains 8,056 observations.

First, the input elasticities of the six most represented industries in the panel are discussed. The industries in case are: *Manufacturing, Construction, Retail trade, Transportation, Professional activities* and *Administrative services.* Tables 6 and 7 include the production function estimation using the LP and FE methods, respectively. Additionally, two-sided Wald tests are performed in order to see whether specific industries exhibit constant returns to scale or not. The labor coefficient appears to range from 0.446 to 0.779, where the highest value is for industries with heavy reliance on labor, such as professional and scientific activities. The estimates are highly significant at the 1% level across all industries. The same observations hold for the capital input. The range of coefficients is wider - from 0.415 for retail and trade to 1.007 for transportation.

Table 6: Levinsohn-Petrin (LP_TFP) estimates for the most represented six industries (2009-2012)

Log of value-added, (VA)						
Industry	Manufacturing	Construction	Retail trade	Transportation	Pro. activities	Admin. services
Classification	Ind. 3	Ind. 6	Ind. 7	Ind. 8	Ind. 13	Ind. 14
	β	β	β	β	β	β
Log of Labor, (L)	0.577***	0.466***	0.462***	0.496***	0.779***	0.446***
	(0.05)	(0.09)	(0.05)	(0.08)	(0.08)	(0.06)
Log of Capital (Assets), (K)	0.646***	0.639***	0.415***	1.007***	0.621***	0.595***
	(0.09)	(0.12)	(0.06)	(0.32)	(0.14)	(0.13)
$L+K$	1.223	1.105	0.877	1.503	1.4	1.041
Wald test for CRS, χ^2 statistic	3.82*	0.75	3.92**	2.43	8.41***	0.07
Constant returns to scale	No	Yes	No	Yes	No	Yes
Observations	1,628	632	2,286	346	864	297

Note: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$ indicate the level of significance. The standard errors of the estimates are reported in parentheses under the β coefficients.

The sum of the coefficients are calculated below the parameter estimates. The Wald test evaluates whether the sum of the two input coefficients are equal to unity, which would indicate constant returns to scale (CRS). This test is applied to all specifications. According to the LP methodology, manufacturing, retail trade and professional services do not display constant returns to scale at a significant level. Even though for the majority of the industries the sum of the coefficients exceeds unity, statistically significant deviation from CRS can only be established for three industries.

Applying the fixed effects estimation with time dummies, produces estimates with higher variation in Table 7. For the manufacturing and construction industries the estimates are robust across the two methodologies and highly similar. However, notable differences in the size of the labor elasticity can be seen for retail and trade, transportation, professional and administrative services. The standard errors have also increased compared to the LP specifications. Only retail and trade and administrative services have either increasing or decreasing returns to scale. For the remaining four industries it can be asserted that they operate under CRS.

The following set of productivity estimates use the Olley-Pakes methodology and compare the results with a Pooled OLS regression. The purpose is to stress the direction of the bias if the endogeneity issues is not handled. The coefficients of the variable input (labor) and semi-fixed input (capital) display the predicted bias. Table 8 show the upward bias in the labor coefficient, $\beta_L^{\hat{\text{OLS}}} > \beta_L^{\hat{\text{OP}}}$ and the capital estimate is biased downwards, $\beta_K^{\hat{\text{OLS}}} < \beta_K^{\hat{\text{OP}}}$.

Table 7: Fixed effects (OLS_TFP) estimates for the most represented six industries (2009-2012)

Log of value-added, (VA)						
Industry	Manufacturing	Construction	Retail trade	Transportation	Pro. activities	Admin. services
Classification	Ind. 3	Ind. 6	Ind. 7	Ind. 8	Ind. 13	Ind. 14
	β	β	β	β	β	β
Log of Labor, (L)	0.553***	0.468***	0.318***	0.388**	0.395***	0.146
	(0.11)	(0.17)	(0.08)	(0.18)	(0.09)	(0.09)
Log of Capital (Assets), (K)	0.329***	0.343***	0.300***	0.467***	0.526***	0.357**
	(0.08)	(0.12)	(0.04)	(0.13)	(0.10)	(0.16)
$L + K$	0.882	0.811	0.618	0.855	0.921	0.503
Wald test for CRS, F statistic	2.41	1.63	20.10***	0.81	0.54	11.59***
Constant returns to scale	Yes	Yes	No	Yes	Yes	No
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,628	632	2,304	347	881	304

Note: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$ indicate the level of significance. The standard errors of the estimates are reported in parentheses under the β coefficients. The constant term has been suppressed from the table output along with the Year FE.

Table 8: Olley and Pakes (OP_TFP) and Pooled OLS estimates for all industrialized companies (2009-2012)

Log of value-added, (VA)		
Estimator	Olley and Pakes (1996)	Pooled OLS
Productivity measure:	OP_TFP	OLS_TFP
	β	β
Log of Labor, (L)	0.417***	0.422***
	(0.02)	(0.01)
Log of Capital (Assets), (K)	0.527***	0.509***
	(0.16)	(0.01)
$L + K$	0.944	0.931
Wald test for CRS, F statistic	24.70***	62.14***
Constant returns to scale:	No	No
Year FE	Yes	Yes
Observations	2,877	6,381

Note: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$ indicate the level of significance. The standard errors of the estimates are robust and are reported in parentheses under the β coefficients. The constant terms have been suppressed from the table output along with the time fixed effects. The main interest is in the coefficients of the two production inputs.

This sub-section presented the input elasticities both by specific industries where possible and for the total sample. Several estimation methods were applied and their results were discussed in reference to each other. The methods produce estimates in a similar range, which is comforting for the robustness across specifications. In the final table the endogeneity issue was underlined with an example of the OP vs. OLS regressions, which identified the bias that occurs if the econometric problem is not taken into account. By using the Index TFP method, the next sub-section delves deep into the productivity analysis of the firms in the sample.

5.2 Comparisons of TFP levels and growth

This sub-section introduces the differences in productivity among multinational firms and those that operate only on the Danish market (no subsidiaries abroad). The key objective is to test whether there are productivity differences between these two types of entities, kernel density estimations are performed along with mean comparisons of productivity across groups and Kolmogorov-Smirnov tests of stochastic dominance. Several TFP measures are used to check the robustness of the findings. The analysis in this section is based on the Index TFP (I_TFP) construct.

Next, the productivity distributions for all 19 industries aggregated together are plotted across time by internationalization status. Several interesting findings are worth mentioning. Figure 4 shows that across the four years the aggregate productivity has shifted to the right for firms engaged in FDI. This suggests that multinationals have been increasing their productivity several years after the financial crisis began. In Figure 4 the productivity distributions for 2009 and 2010 almost perfectly coincide, whereas from 2011 a notable shift can be seen. What the two figures do not show are the individual TFPs of firms over the time period. If the Index TFP measure was constructed without the chain-linking components it would have not reflected correctly the actual development in efficiency among firms.[15]

Figure 4: Chain-linked Index TFP development of MNEs, 2009-2012

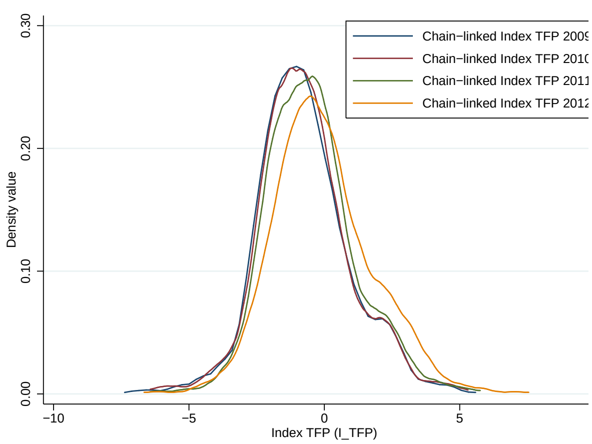


Figure 5: Chain-linked Index TFP development of purely domestic firms, 2009-2012

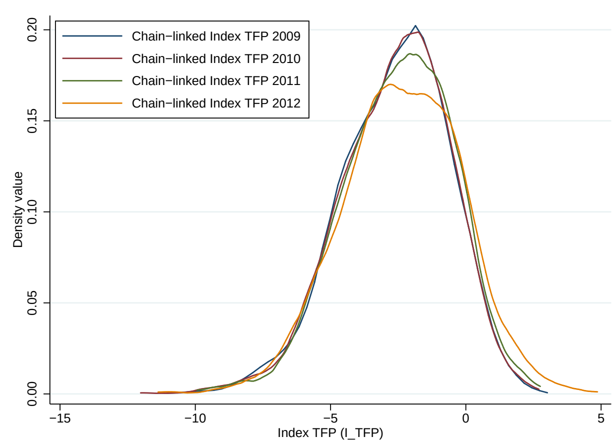


Figure 5 shows that the efficiency of purely domestic firms has not improved drastically as the shift towards the right is almost unnoticeable. Some density around the mean is lost in favor of a widening distribution and gain of mass close to the tails. The distribution of multinationals' productivity has much smaller dispersion compared to the purely domestic companies, which is consistent with previous findings such as those of Aw, Chung and Roberts (2000) and Girma, Kneller and Pisu (2005). Tighter competition and higher productivity in order to operate on international markets, make firms converge towards a higher average productivity level, which can be seen from the tails of the distributions in

[15] In order to illustrate the impact of the two chain-linking parts of the multilateral index, Appendix section D.1 illustrates two figures, which show the shift in productivity distribution. For multinational companies the impact is larger than for purely domestic firms because shift of the distribution in the former case is larger than for the latter. In the multilateral index, the year 2009 is used as a base year and the figures in the Appendix are showing the differences in efficiency for 2012.

the Figures 4 and 5. Multinationals have a dispersion of the multilateral index in the range: -7 to 7.55 and purely local companies in the band: -11.38 to 4.86 over the years. The wider dispersion and slower shift toward higher efficiency of purely domestic firms can be confirmed by looking at Table 9. Over the four years, the efficiency mean for multinationals has increased towards higher productivity by 0.698 whereas for domestically operating companies this increase is only 0.252. This observation suggests that firms operating internationally have experienced a bigger aggregate TFP shift than non-multinationals. In 2012 there is more convergence towards a higher mean productivity compared to 2009 of MNEs with subsidiaries across regions. It can also be seen that the 25th and 75th percentile values are much closer to the average TFP in the case of firms engaged with FDI compared to non-multinationals.

Table 9: Descriptive statistics for Index TFP of firms by year and international status

Multinationals	2009	2010	2011	2012	Non-multinationals	2009	2010	2011	2012
Mean	-0.748	-0.707	-0.470	-0.050	*Mean*	-2.618	-2.616	-2.494	-2.366
Std. dev.	1.672	1.685	1.656	1.844	*Std. dev.*	1.985	2.013	2.028	2.214
Max.	5.210	5.305	5.757	7.549	*Max.*	2.596	2.704	2.752	4.866
Min.	-7.000	-6.428	-6.436	-6.657	*Min.*	-10.619	-12.023	-9.933	-11.375
25th percentile	-1.856	-1.815	-1.597	-1.244	*25th percentile*	-3.981	-3.967	-3.841	-3.801
75th percentile	0.157	0.191	0.428	0.970	*75th percentile*	-1.200	-1.184	-1.011	-0.782
N	580	603	613	617	*N*	1,401	1,383	1,380	1,379

Figures 6 and 7 plot the efficiency density for purely domestic companies and multinationals doing FDI in specific regions of the world for 2009 and 2012, respectively. Since 2009 is the base year, it can be stated that a big dispersion can be observed among companies doing FDI in different areas compared to the much more similar efficiency distributions of MNEs in 2012. The further segmentation of FDI doers into seven separate regions is important, since there exists TFP heterogeneity even among multinationals. The paper is novel in the sense that it takes a detailed look at firms involved in FDI at a granular level. From the two figures it becomes clear that productivity distribution of non-MNEs is further to the left in relation to all the efficiency distributions of MNEs. The results are further supported by the statistical tests performed next.

The next step is to assess the statistical significance of the productivity distributions between the two main types of firms in the sample. In a second step, the productivity levels of firms with FDI in specific geographical regions are benchmarked against non-multinational companies in an attempt to obtain a strict ranking of productivity. This test is very useful as it goes beyond the comparison of means that was performed for Table 9.

Following Girma, Kneller and Pisu (2005), non-parametric Kolmogorov-Smirnov tests of stochastic dominance are applied for each cross-section of TFP observations by internationalization status. Table 10 offers the results of the K-S test, applied on each cross-section of observations. In each row the TFP distribution of the purely domestic companies (non-MNEs) is assessed against MNEs engaged in FDI in a specific region. In the first row of Table 10 the null hypothesis of equality of TFP distributions is rejected at the 1% level, supporting the statement that companies having subsidiaries in Europe are more productive than their counterparts that only conduct business in Denmark. This finding

Figure 6: Chain-linked Index TFP development of MNEs with subsidiaries across regions, 2009

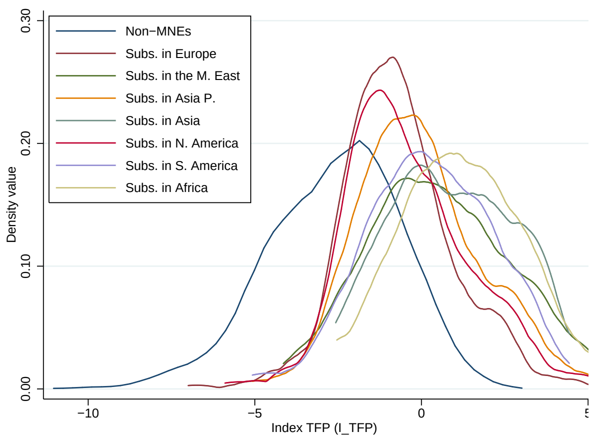
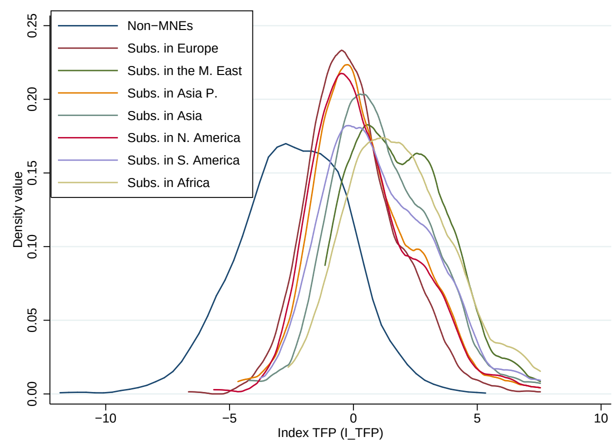


Figure 7: Chain-linked Index TFP development of MNEs with subsidiaries across regions, 2012



is consistent across time. Similarly, there are significant differences in the distribution of the change in TFP between non-MNEs and MNEs having activities in Europe. The stochastic dominance of the distribution of MNEs (EU) means that on average multinationals are growing faster than non-multinationals. All FDI doers have TFP distributions that dominate the efficiency of non-MNEs. For four pairs, the growth of TFP, observed in MNEs is always dominating the efficiency progress for purely domestic companies for all the four years. Only in 2009, the change in TFP is not significantly different between domestic companies and those with subsidiaries in Asia, Asia Pacific and Africa. After 2009, notable differences in the distribution of both the multilateral TFP index and the change in TFP can be concluded.

Table 10: Kolmgorov-Smirnov test for Non-MNEs vs. MNEs with subsidiaries by region and year

Pair	I_TFP - Index TFP				d_I_TFP - Change in Index TFP		
	2009	2010	2011	2012	2010	2011	2012
Non-MNEs and MNEs (EU)							
K-S stat.	0.401***	0.411***	0.423***	0.453***	0.110***	0.176***	0.259***
Difference in TFP	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Non-MNEs and MNEs (ME)							
K-S stat.	0.574***	0.705***	0.673***	0.732***	0.214**	0.267***	0.420***
Difference in TFP	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Non-MNEs and MNEs (ASP)							
K-S stat.	0.497***	0.493***	0.490***	0.528***	0.103*	0.151***	0.310***
Difference in TFP	Yes	Yes	Yes	Yes	No	Yes	Yes
Non-MNEs and MNEs (AS)							
K-S stat.	0.623***	0.615***	0.610***	0.634***	0.139*	0.185***	0.329***
Difference in TFP	Yes	Yes	Yes	Yes	No	Yes	Yes
Non-MNEs and MNEs (NA)							
K-S stat.	0.444***	0.467***	0.457***	0.505***	0.136***	0.209***	0.325***
Difference in TFP	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Non-MNEs and MNEs (SA)							
K-S stat.	0.566***	0.593***	0.555***	0.568***	0.242***	0.234***	0.302***
Difference in TFP	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Non-MNEs and MNEs (AF)							
K-S stat.	0.736***	0.731***	0.711***	0.712***	0.132	0.250***	0.336***
Difference in TFP	Yes	Yes	Yes	Yes	No	Yes	Yes

Note: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$ indicate the level of significance of the K-S statistic, which is the maximum difference between the two tested distributions. The difference in TFP levels or change (Yes/No) are evaluated at a 95% confidence level, meaning that if the p-value exceeds 0.05, then it is considered that there are is *no* difference in TFP levels or change.

To supplement the K-S tests, Figures 8 and 9 plot the distribution of TFP growth for 2010 and 2012, respectively. Upon closer inspection, on average MNEs have had higher productivity growth compared to purely domestic companies. The dark blue density is to the left compared to the other distributions. This difference becomes even more pronounced in 2012 where all MNEs display noticeably higher growth rates on average than their counterparts solely operating in Denmark.

Table 11 shows that the mean growth has been increasing for both types of companies, but the increase within these four years has been much larger for MNEs than for domestically selling companies.[16] These observations are in line with previous research, which discovered that MNEs are more productive and grow faster in terms of efficiency (Girma, Kneller and Pisu, 2005; Manova and Zhang, 2012). The overall conclusion from the analyses of the aggregate productivity is that there are significant differences in productivity levels and growth rates between non multinationals and firms operating on the international markets.

[16] The difference in mean TFP change for multinationals between 2009 and 2012 is $0.441 - 0.085 = 0.356$, whereas for non-MNEs it is only $0.138 - 0.026 = 0.112$.

Figure 8: Change in chain-linked Index TFP of MNEs with subsidiaries across regions, 2010

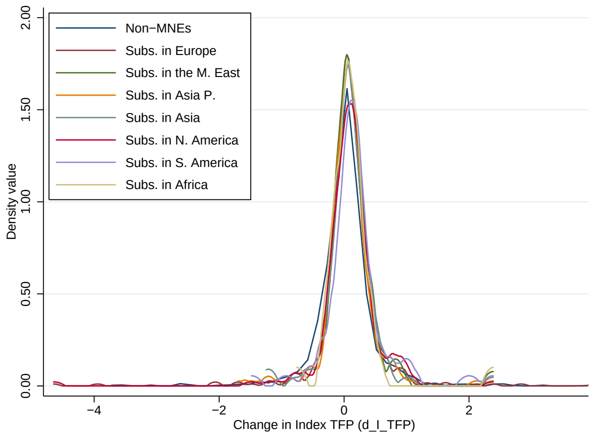


Figure 9: Change in chain-linked Index TFP of MNEs with subsidiaries across regions, 2012

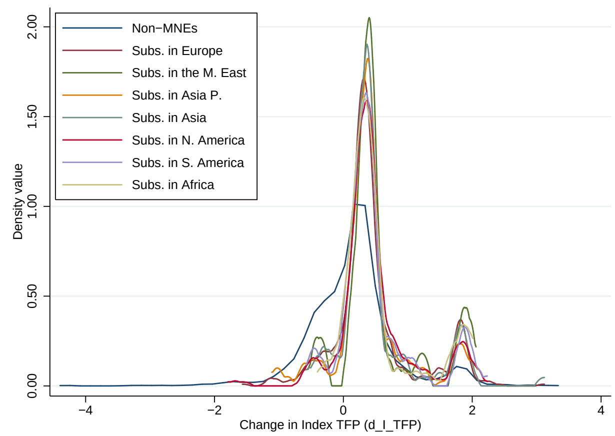


Table 11: Comparative statistics for the change in Index TFP of firms by year and international status

Multinationals	2010	2011	2012	Non-multinationals	2010	2011	2012
Mean	0.085	0.256	0.441	*Mean*	0.026	0.120	0.138
Std. dev.	0.511	0.447	0.632	*Std. dev.*	0.493	0.488	0.680
Max.	4.017	3.154	3.121	*Max.*	3.917	4.210	3.238
Min.	-4.649	-2.071	-1.793	*Min.*	-3.666	-3.227	-4.300
25th percentile	-0.067	0.057	0.181	*25th percentile*	-0.159	-0.102	-0.228
75th percentile	0.250	0.395	0.511	*75th percentile*	0.191	0.305	0.389
N	600	607	614	*N*	1,359	1,368	1,368

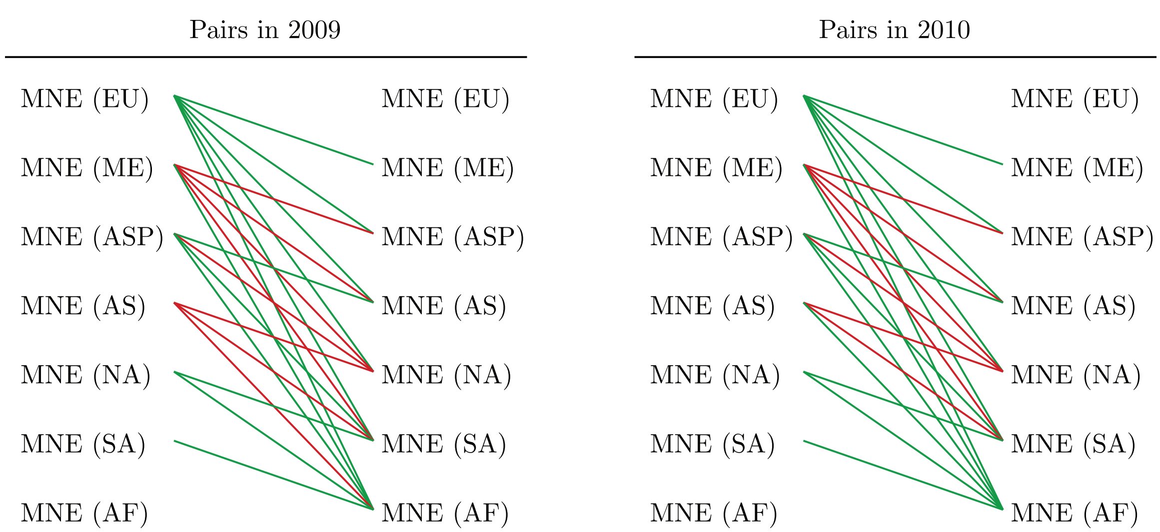
So far the paper has established the TFP differences on an aggregate level between multinationals and purely domestic companies by treating MNEs as a single entity. The second set of research questions aims at understanding whether there are any differences and ranking *among* multinational firms. These questions are tackled sub-section 5.3, which analyzes the differences in levels and growth rates of TFP of multinational Danish companies with respect to their choice of FDI host region. This step is a novel contribution as the analysis tries to uncover well established differences among firms that are already serving foreign markets. In that sense, particular regions could be ranked with respect to the average efficiency need to operate there.

5.3 Cross-regional productivity comparison

This section takes the analysis a step further as it applies the K-S tests on the TFP distributions of each pair of FDI locations. This allows to the answer the questions whether firms that do FDI in a given region grow faster or have higher mean TFP levels compared to firms that have international presence in another reference region. In total seven regions are tested against each other by means of Kolmogorov-Smirnov tests and graphical analysis. The kernel densities of the TFPs and the change in TFP were presented in Figure 6 through Figure 9 in the previous sub-section.

In each pair, the hypothesis that there are no significant differences between the productivity distributions is tested on a yearly basis. The results of the detailed Kolmogorov-Smirnov tests are available in the Appendix section D.1. However, a novel graphic approach is presented in Figures 10 and 11. First, in Figure 10, the green connections mean that the null hypothesis of the two-sided K-S test is rejected in favor of the alternative that there evaluated TFP distributions are significantly different. In contract, the red connecting lines suggest that the null hypothesis cannot be rejected and the two tested distributions of productivity are not different from each other at a standard 5% significance level.

Figure 10: Differences in the Index TFP distributions of multinationals with FDI by regions, 2009-2010

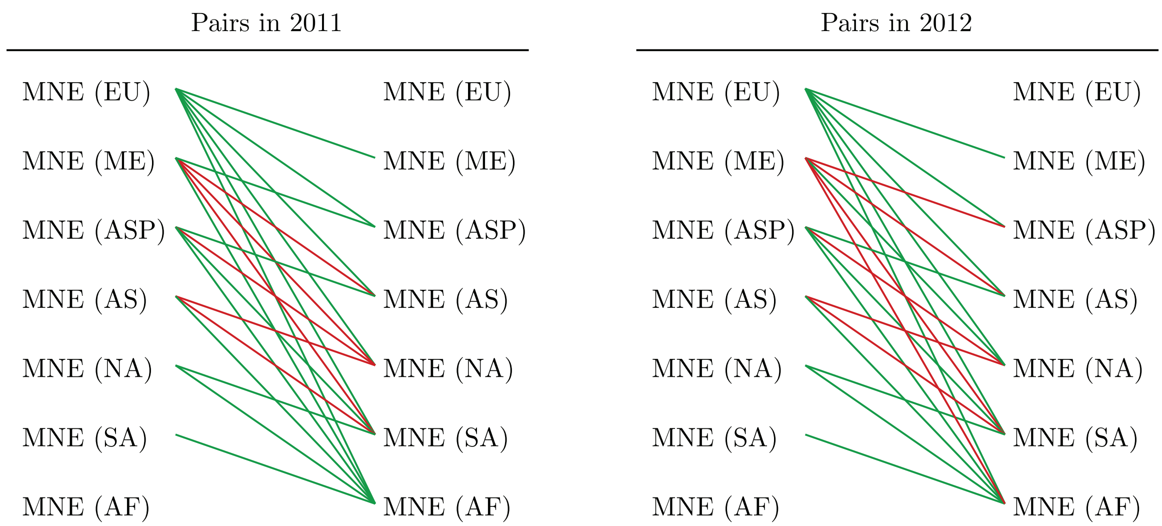


Note: A green link indicates significant differences between the two tested distributions at a 5% level while a red link indicates that the distributions do not differ statistically.

For 2009 and 2010 the patterns are very similar. Firms with FDI in Europe have significantly different productivity levels compared to all other regions, as it can be concluded from the green links. In 2009 firms with FDI in Asia did not have different productivity levels than firms doing FDI in North and South America and in Africa. In the following year there were significant differences between Danish multinationals with FDI in Africa and Europe, as indicated by the green line in Figure 10.

Figure 11 shows that in 2012 companies with subsidiaries in the Middle East do not have stochastically dominating TFP levels compared to firms doing FDI in other regions besides Europe and North America. Since the number of pairs in which there are no differences between the efficiency levels, these pairs are stated next. In the final year of observation there were no significant efficiency differences between firms with FDI in: ME and ASP, ME and AS, ME and SA, ME and AF. The list is supplemented with the following pairs: ASP and NA, AS and NA, AS and SA.

Figure 11: Differences in the Index TFP distributions of multinationals with FDI by regions, 2011-2012

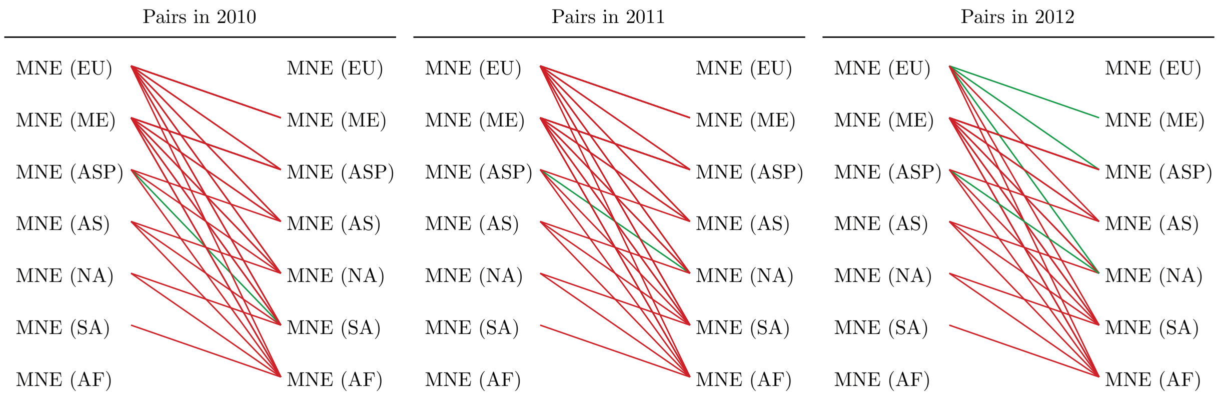


Note: A green link indicates significant differences between the two tested distributions at a 5% level while a red link indicates that the distributions do not differ statistically.

The same approach can be applied to the remaining regions. In 2009 the number of pairs in which the null hypothesis could not be rejected is eight, in 2010 - seven, in 2011 - six and in 2012 - seven. These findings suggest that overall there has been a trend towards a widening of productivity among multinationals, supported by the observation that the number of pairs that exhibit similar productivity distributions have been declining with a small increase in the final year. This slightly bigger spread can be visually confirmed by looking at the scale and shape of the efficiency density functions in Figures 6 and 7.

The other constituent part of the pairwise productivity analysis is to look at the TFP change and evaluate the distributions. The complete results of the K-S tests for the cross-regional comparisons are presented in Tables 26, 27, 28 and 29 in Appendix section D.1. Figure 12 contains the graphical output of those tests. The predominantly red links suggest that in almost all the cases the null hypothesis cannot be rejected of the K-S tests. This means that statistically the distribution of TFP growth does not differ significantly within different host FDI pairs. In 2010 the change in TFP of firms doing FDI in Asia Pacific differed from that of firms doing FDI in South America. In the following year, the only significant difference in the distributions is between Asia Pacific and North America. The change development of TFP for the last year reveals that growth rates have diverging for firms doing FDI in Europe vis-à-vis companies, which doing FDI in the Middle East, Asia Pacific and North America. The next sub-section ranks the host FDI locations with respect to the average productivity of the firms.

Figure 12: Differences in distributions of Index TFP growth of multinationals with FDI by regions, 2010-2012



Note: A green link indicates significant differences between the two tested distributions at a 5% level while a red link indicates that the distributions do not differ statistically.

5.4 Productivity ordering

The cross-pair tests on the productivity distributions of firms doing FDI abroad showed that in terms of growth, there are almost no significant differences in the TFP changes from year to year across destinations, besides in 2012. However, the TFP distributions vary significantly across regions as suggested by Figures 10 and 11. All MNEs differ substantially from non-multinationals both in terms of productivity level but also in terms of growth dynamics. Confirming the insights from the kernel estimates, Danish companies that had FDI in the Middle East demonstrate one of the highest changes in TFP in 2010, trailing only behind firms with subsidiaries in South America. The popularity of Asia has been increasing in the last several years as it can be seen from the average changes in TFP for the sample period.

These result suggest that firms investing in Asia have been increasing higher change in their TFP, compared to firms doing FDI in Europe and Asia Pacific. In 2012, the average change in productivity has been highest for firms doing FDI in geographical regions such as the Middle East, Africa and Asia. These patterns are not surprising as individual countries in these area benefit from economic development. In turn this attracts foreign direct investments both due to cheaper labor force but also because of the enlarging domestic demand in those places.

It is important to state that average changes in efficiency are important but it is also crucial to understand the initial distribution of TFP. Some regions may have higher growth in average productivity of firms that do FDI there, but they could initially start from a much lower average productivity. From the Table 12 no clear pattern appears to be visible in the average changes in TFP. This fully supports the K-S tests, visually represented in Figure 12, where no clear differences were identified in the TFP growth distributions among pairs. Similar results are obtained in Girma, Kneller and Pisu (2005) who obtain rankings of non-MNEs, exporters and firms engaged in FDI. In between the four years of observations the productivity ranking of FDI host destinations has not changed drastically with the only exception of North and South America, which did not preserve their initial ranking from 2009. Overall, firms that do have subsidiaries in the Middle East and Africa show the highest average

TFP levels, followed by Asia, South America, Asia Pacific, North American and Europe.

These results have to be analyzed with utmost care. This ranking does not mean that the most productive Danish multinationals choose the Middle East and Africa as their preferred locations to establish subsidiaries, but rather that only *the most productive* firms can afford to participate on these markets and establish foreign affiliates by self-selecting into them. Geographical areas such as Europe, North America and Asia Pacific are among the FDI destinations for which multinational firms have the lowest average TFP. The reason is that establishing a foreign subsidiary is much easier and less costly compared to other destinations. The common cultural ties in Europe, North America and the countries from the Common Wealth (belonging to Asia Pacific) are easier to be accessed, compared to the remaining locations. In other words, the mentioned productivity-host FDI region can be interpreted as a scale, which shows how easy it is to serve the given geographical area by FDI. Relating to the profit functions of multinational firms in Appendix A, the fixed costs of doing FDI (f_{FDI}) in the various regions are different. This means that f_{FDI} can be strictly ranked by location based on the obtained results from the kernel estimates and the K-S tests of stochastic dominance.

Table 12: Mean productivity levels and change in Index TFP by year and region of FDI

	Mean levels of Index TFP						
	MNE (EU)	MNE (ME)	MNE (ASP)	MNE (AS)	MNE (NA)	MNE (SA)	MNE (AF)
2009	-0.692	0.498	-0.068	0.883	-0.400	0.265	1.107
2010	-0.659	0.911	-0.060	0.460	-0.232	0.472	1.316
2011	-0.434	1.010	0.096	0.662	-0.082	0.615	1.346
2012	0.024	1.878	0.481	1.133	0.417	1.037	1.893

	Mean change in Index TFP						
	MNE (EU)	MNE (ME)	MNE (ASP)	MNE (AS)	MNE (NA)	MNE (SA)	MNE (AF)
2010	0.083	0.177	0.078	0.091	0.084	0.184	0.113
2011	0.259	0.208	0.193	0.233	0.234	0.227	0.207
2012	0.453	0.571	0.429	0.503	0.481	0.490	0.551

According to the 2012 ranking from Table 12, the ranking of fixed costs looks as follows: $f_{FDI}^{AF} > f_{FDI}^{ME} > f_{FDI}^{AS} > f_{FDI}^{SA} > f_{FDI}^{ASP} > f_{FDI}^{NA} > f_{FDI}^{EU}$. The productivity differentials among host destinations also support the pattern related to the number of firms in each region. Table 4 shows that most firms in the sample prefer Europe as a FDI host destination. Very few Danish companies have managed to establish affiliates operations in regions with weaker business ties, without a common language, smaller markets and less attractive investment conditions. The ranking pattern is in line with other studies such as Nunnenkamp and Spatz (2004) and Business (2011), emphasizing the importance of the host country characteristics in attracting FDI flows. The ranking of the business friendliness environment fully coincides with the ranking in the costs associated with FDI by region. The theoretical model from section A in Appendix suggests that in order to serve those markets by FDI, a much higher productivity is necessary to recuperate the costs. The productivity ranking of Danish MNEs by FDI host regions is a focal contribution. It shows that the majority of firms set up affiliates in regions that are easier to be served and to do business in. The empirical results suggest that MNEs shall not be treated as a unified group but be studies at a further level of disaggregation.

5.5 Learning-by-doing FDI or self-selection?

The suggested mechanism in Navaretti and Castellani (2004) is that firms adopt best practices in their fields after they have set up a foreign affiliate, which raises their productivity level. On the other hand, simply better firms may self-select into doing FDI as they are already more productive even before becoming multinationals. The theoretical motivation for firms doing FDI in the models of Melitz (2003) and Helpman, Melitz and Yeaple (2004) is self-selection. The dynamics of the model suggest that companies that are the most productive in the population of firms decide to establish foreign subsidiaries because only they are able to recuperate the higher fixed costs by exhibiting higher productivity.

One way the learning-by-doing FDI hypothesis could be investigated is by splitting the sample firms into entrants, continuers, exiters and purely domestic companies that never had any foreign affiliates with the studies time frame. Entrants are defined as firms that in 2009 had no subsidiaries but in later years established a foreign affiliate; Continuers are Danish multinationals that throughout the sample period 2009-2012 had at least one subsidiary; Exiters are firms that in 2009 had a foreign affiliate but in 2012 did not; Domestic companies are firms that in the period 2009-2012 never undertook FDI. If the average growth of TFP for entrants exceeds that of continuers, exiters and purely domestic firms, then it can be stated that firms learn by doing FDI abroad, which would lead to convergence with the more experienced companies. However if there are no significant differences in the TFP growth distributions, this would mean that there is no convergence.

The Kolmgorov-Smirnov tests compare the TFP distributions of entrants, continuers and purely domestic companies from 2009 and 2012. Additional analysis is performed on the change in TFP distributions between the two categories. The results in Table 13 reject the learning-by-doing FDI hypothesis. The reason is that there no statistical differences between the TFP growth distributions between entrants and continuers, meaning that entrants are not converging to the higher level of efficiency of firms already doing FDI abroad. Furthermore, in 2012 the productivity distributions of continuers dominates the efficiency distribution of entrants. In comparison to purely domestic companies, entrants show significantly higher growth in 2012. In 2009 the growth distributions did not differ significantly from each other. Across all the four years, the TFP levels of entering companies have been stochastically dominating the efficiency of purely domestic companies. The same applies to continuers who experience both higher TFP levels but also growth relative to purely domestic companies as suggested by the last row of Table 13. Having in mind the small number of observations for two of the four categories, the robustness of these results shall be tested by employing longer time series.[17]

[17]Purely domestic firms, entrants, exiters and continuers represent 68.12%, 1.84%, 1.44% and 28.60% of the sample, respectively. The low number of entrants could be explained by the short time frame that is analyzed but also by the fact that after the financial crisis, industrialized firms have revised their international expansion prospects.

Table 13: Kolmgorov-Smirnov test for entrants, continuers and purely domestic firms by year

Pair	Index TFP		Change in Index TFP	
	2009	2012	2010	2012
Purely domestic firms and Entrants				
K-S stat.	0.337**	0.401***	0.118	0.233**
Difference in distributions	Yes	Yes	No	Yes
Entrants and Continuers				
K-S stat.	0.216*	0.238**	0.200	0.178
Difference in distributions	No	Yes	No	No
Purely domestic firms and Continuers				
K-S stat.	0.403***	0.461***	0.112***	0.264***
Difference in distributions	Yes	Yes	Yes	Yes

Note: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$ indicate the level of significance of the K-S statistic, which is the maximum difference between the two tested distributions. The difference in the TFP and change in TFP distributions (Yes/No) are evaluated at a 95% confidence level, meaning that if the p-value exceeds 0.05, then it is considered that there are is *no* difference in distributions.

5.6 FDI productivity premium

This sub-section quantifies the impac on productivity of doing FDI compared to only doing business domestically. The analysis that has been conducted so far only acknowledges that there are significant differences in productivity among MNEs and purely domestic firms, but it does not answer how big the advantage is and whether it is significant. By specifying an econometric model, one can account for industry heterogeneity, firms size and other observable factors that if not taken into account would blur the true FDI premium. Assuming that the model assumptions hold and there is no omitted variable bias, the pure premium from being a MNE can be estimated. The term "productivity premium" is used often in the empirical trade literature. For example studies such as Kimura and Kiyota (2006); De Loecker (2007); Delgado, Farinas and Ruano (2002) investigate exporter premiums. Due to the data limitations, the analysis is restricted only to the FDI premium.

Both growth and levels of Index TFP are regressed on a vector of independent variables, which have been shown in the literature to affect productivity. Host FDI regional dummies are added in order to identify the differences in premiums experienced by firms. The first specification takes the following form:

$$\theta_{ijt} = \alpha_0 + \beta_1 \mathbf{X}_{it} + \beta_2 \mathbf{K}_{ijt} + \sum_{t=2009}^{2012} \beta_t Year_t + \sum_{j=1}^{19} \beta_j Industry_j + u_{ijt} \text{ for } t = 2009, ..., 2012 \quad (6)$$

, where θ_{it} is the Index TFP measure for firm i at time t, $\mathbf{X}_{it}$ is a vector of dummies identifying host FDI regions in which firm i at time t had subsidiaries. Another version of Equation 6 incorporates a dummy variable, called $status$, which is equals 1 (MNE) if a given company has at least one subsidiary and 0 (Purely domestic) otherwise. It is a more aggregated variable as it does not differentiate the location of the subsidiaries. Vector $\mathbf{K}_{ijt}$ contains the firm specific factors, such as log of number of employees and the age of the company in years. The reason why these independent variables can be included is that usually bigger firms with more employees have more subsidiaries abroad (Head and Ries, 2003). The same logic applies to the age of firms, which would mean that older firms are expected to be more experienced in their operations, leading to higher productivity. The four $Year_t$ and 19 $Industry_j$ dummies are included in the model to capture time effects and industry and firm heterogeneity. To avoid perfect collinearity issues, for each of the categorical variables there is a reference group, which is left out of the equation. Finally, u_{ijt} is an error term.

The second model, estimates the change in TFP. The estimation window is reduced for the period 2010-2012. The purpose is to see whether doing FDI accelerates the productivity of multinationals. The estimated equation looks as follows:

$$\Delta\theta_{ijt} = \alpha_0 + \beta_1\mathbf{X}_{it} + \beta_2\mathbf{K}_{ijt} + \sum_{t=2010}^{2012}\beta_t Year_t + \sum_{j=1}^{19}\beta_j Industry_j + u_{ijt} \text{ for } t = 2010, ..., 2012 \quad (7)$$

, where the explanatory variables are the same but the dependent variable, $\Delta\theta_{ijt} = \theta_{ijt} - \theta_{ij,t-1}$ is the year-to-year change in productivity. The dataset at hand includes the number of subsidiaries and status of each firm for all the year, which means that more time periods can be considered in the estimation. Similar empirical studies estimated a cross-sectional regression of efficiency growth by only regression the difference in productivity between the initial and terminal period. Adopting this approach would not allow the inclusion of time effects and dynamics in the growth pattern across time. Equation 7 is preferred than the described alternative as it takes into account the temporal developments in productivity and not only the cross-sectional nature of the data.

The benefit of the Fixed effects (FE) method is that allows to capture all time-invariant unobserved heterogeneity through data demeaning. All factors that affect the productivity of company, which are not included in the model and are not varying across time are wiped away due to the FE transformation. There are several disadvantages of using this approach when quantifying FDI premiums. The problem is that the FE estimator takes only the *within* panel variation when estimating the coefficients in Equations 6 and 7. So firms that change their status within the four years of observation are taken into account. For example, firms that are continuers or only operating domestically (Non-MNEs) are not included in the estimation of the α_0 and β coefficients. Only entrants and exiters change their internationalization status within their panels.[18]

In contrast to the FE estimation, Random effects (RE) applies a GLS estimator and takes advantage of the *between* variation in panels in order to estimate the β coefficients. It makes use of all the observation as the time-constant variables are not wiped away, since the data is quasi-demeaned. However, strict assumptions apply on the relationship between the unobserved time-invariant factors and the regression covariates. A Hausman test was performed in order to assess whether FE or RE is the more appropriate method. The results are mixed as in some model specifications, RE was preferred and in others - FE. This is why it is decided to apply both methods for comparison purposes. The necessary assumptions of the two methods is discussed in sub-section **??**. The following sub-section presents the final results of the empirical models in Equations 6 and 7.

[18]The dummies for continuers and purely domestic firms were defined in sub-section 5.5.

5.7 FDI premium results

First, the results for the levels of TFP and FDI premiums are presented. In the initial specification there is no clear differentiation between the location of subsidiaries. Only the binary variable *status* is included, which identifies MNEs with 1 if a firm has at least one subsidiary and 0 otherwise, without explicit reference to the regions of FDI. Afterwards, a full set of results is presented for each host region.

Table 14 offers interesting insights about the size of the FDI premium by specifying several models and estimating them with three different methods. The first specification in each column includes firms characteristics (age and the logarithm of number of employees), year and industry dummies, accompanied by a dummy variable named "FDI premium" (in the dataset it is coded under the name "*status*"). The second equation in each column only regresses productivity on the FDI dummy, without accounting for firm-specific factors. The third specification in each column includes all firm characteristic, but instead of using the aggregate level dummy "FDI premium" it features individual dummies for every FDI region where Danish firms have subsidiaries. In that way a FDI premium can be estimated by host region. The fourth specification is the same as the third except it does not include the firm characteristics. The results from the alternative productivity measures are highly similar to the main coefficients offered in the main part of the analysis section. The labor TFP estimates are present in Tables 30 and 31 in section D.1. The inherent econometric issues of the estimated models are offered in and **??** in the Appendix.

The results for the Pooled OLS estimation suggest that there is a significant FDI premium for MNEs compared to domestic firms. These results are supported by the K-S tests in the previous sub-sections. Since the relationship between TFP and the "*status*" dummy in the model is log-level, it means that in order to get the percentage impact on productivity, the true percentage FDI premium is calculated as $100 \times (\exp^{\beta} - 1)$.

Table 14: FDI premiums - POLS, FE and RE estimates of Index TFP levels

θ	Pooled OLS				Fixed Effects				Random Effects			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
FDI premium	0.175***	2.119***			0.024	0.043			0.152***	0.611***		
	(0.03)	(0.05)			(0.07)	(0.06)			(0.04)	(0.07)		
Age	0.002***		0.002***		0.148***		0.147***		0.003***		0.003***	
	(0.00)		(0.00)		(0.01)		(0.01)		(0.00)		(0.00)	
Log employees	0.970***		0.969***		0.579***		0.575***		0.944***		0.941***	
	(0.01)		(0.01)		(0.05)		(0.05)		(0.01)		(0.01)	
MNE (EU)			0.095***	1.646***			-0.020	-0.003			0.073***	0.476***
			(0.03)	(0.05)			(0.06)	(0.06)			(0.04)	(0.06)
MNE (ME)			0.006	0.602***			0.103	0.144			0.074	0.278***
			(0.08)	(0.13)			(0.08)	(0.09)			(0.07)	(0.09)
MNE (ASP)			0.011	0.479***			-0.016	-0.035			0.017	0.139 **
			(0.05)	(0.08)			(0.06)	(0.06)			(0.05)	(0.07)
MNE (AS)			0.114*	0.964***			0.148**	0.176 **			0.144**	0.330***
			(0.06)	(0.11)			(0.07)	(0.07)			(0.06)	(0.07)
MNE (NA)			-0.007	0.355***			0.071	0.096			0.041	0.331***
			(0.04)	(0.08)			(0.08)	(0.08)			(0.06)	(0.08)
MNE (SA)			0.130	0.137			-0.020	-0.037			0.060	0.110
			(0.08)	(0.12)			(0.10)	(0.11)			(0.09)	(0.11)
MNE (AF)			0.246***	1.134***			0.092	0.084			0.195**	0.409***
			(0.08)	(0.15)			(0.12)	(0.13)			(0.09)	(0.14)
R^2	0.833	0.282	0.833	0.312	0.265	0.016	0.267	0.032	0.833	0.205	0.828	0.241
Observations:	7,956	7,956	7,956	7,956	7,956	7,956	7,956	7,956	7,956	7,956	7,956	7,956
Firm specifics:	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No

Note: *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$ indicate the level of significance. Robust standard errors are used in the estimations. In the POLS column, the errors have been clustered by firm ID. The constant terms have been suppressed from the table output along with the year and industry dummy variables.

The estimated significant FDI premiums across the different models range from 15.4% to 19% for Pooled OLS and RE in (1). Only the estimates from the POLS and RE estimations are significant at the 1% level. This is totally in line with the K-S tests, which indicated that there are big differences in productivity distributions of MNEs compared to domestic firms. The R^2 is much higher for the complete specifications in (1) than in (2).

In the econometric model without firm characteristics the explained variance is much smaller, indicating that the firm-specific factors have a highly explanatory influence on TFP. The specifications with the individual regional FDI dummies are rarely significant. It is not possible to confirm the previously observed FDI premium for all regions after controlling for firm characteristics. However, across all specifications doing FDI in Asia is linked with a significant increase in productivity. Such a link can also be seen for companies doing FDI in Europe and Africa. An important caveat in those estimates is that the number of firms with subsidiaries in some regions is very small. The number of observations for each category may not be enough to establish statistical significance. Increasing the sample and running the models again may be a desirable exercise. For that reason, the main focus should be on the aggregate dummy - FDI Premium ("*status*") in specifications (1) and (2).

Next, Table 15 features the results from the estimation of Equation 7. Overall, the results are inline with the K-S tests. Looking first at the FDI premium and the specifications in (1) and (2) it can be seen that only the Pooled OLS and RE estimates are significant in the specifications without firm characteristics. It is estimated that firms that are MNEs enjoy around 3.2% premium in their TFP growth compared to non-MNEs. However, for the estimates in (3) and (4) support for significantly different growth pattern is not found. Only Danish firms with FDI in Europe enjoy positive growth premiums at a 5% significance level. However, once it is controlled for firm characteristics the FDI premium in growth disappears. This is consistent with the previous findings where it was asserted that MNEs and non-MNEs do not grow at different rates, but do exhibit different productivity levels.

The explanatory power of the model for TFP change is much smaller and is around 0.286 across all specifications. Furthermore, the number of observations are lower because in the growth computation 2009 is excluded from the regressions. For the two econometric exercises, the overall results are in line with the previous findings. However, the important contribution of this sub-section is that evaluates the productivity differences between MNEs and non-MNEs by controlling for age and size of the company (proxied by the log of the number of employees). In contrast to the TFP distribution tests, where it was found that each FDI region has an efficiency distribution to the right of purely domestic firms, the econometric analysis shows that after controlling for firm-characteristics, the FDI premium for some areas disappears. The estimates of TFP growth in Table 15 show that when firm characteristics are controlled for, the growth premium of almost all FDI host regions disappears. The only exceptions are North America and Asia Pacific that retrain the significance of their estimates even after controlling for other variables in specifications (3) and (4).

Table 15: FDI premiums - POLS, FE and RE estimates of Index TFP change

$\Delta\theta$	Pooled OLS				Fixed Effects				Random Effects			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
FDI premium	0.001	0.032^{**}			0.015	0.026			0.001	0.032^{***}		
	(0.02)	(0.01)			(0.07)	(0.07)			(0.01)	(0.01)		
Age	0.002		0.001		0.089^{***}		0.089^{***}		0.001		0.001	
	(0.00)		(0.00)		(0.01)		(0.01)		(0.00)		(0.00)	
Log employees	0.017^{***}		0.017^{***}		0.356^{***}		0.352^{***}		0.017^{***}		0.017^{***}	
	(0.00)		(0.00)		(0.07)		(0.07)		(0.00)		(0.00)	
MNE (EU)			0.005	0.031^{*}			0.044	0.053			0.005	0.031^{***}
			(0.02)	(0.02)			(0.08)	(0.08)			(0.01)	(0.01)
MNE (ME)			0.008	0.019			0.061	0.087			0.008	0.019
			(0.04)	(0.04)			(0.11)	(0.12)			(0.04)	(0.04)
MNE (ASP)			-0.062^{**}	-0.054^{*}			-0.167	-0.177			-0.062^{**}	-0.054^{**}
			(0.03)	(0.03)			(0.13)	(0.13)			(0.02)	(0.02)
MNE (AS)			0.003	0.016			0.020	0.048			0.003	0.016
			(0.03)	(0.03)			(0.09)	(0.10)			(0.02)	(0.02)
MNE (NA)			0.036	0.042			-0.004	0.008			0.036^{*}	0.042^{**}
			(0.03)	(0.03)			(0.08)	(0.08)			(0.02)	(0.02)
MNE (SA)			0.026	0.026			-0.081	-0.120			0.026	0.026
			(0.04)	(0.04)			(0.11)	(0.12)			(0.03)	(0.03)
MNE (AF)			-0.029	-0.015			0.112	0.137			-0.029	-0.015
			(0.04)	(0.04)			(0.09)	(0.10)			(0.03)	(0.03)
R^2	0.288	0.286	0.289	0.286	0.044	0.029	0.045	0.031	0.287	0.286	0.289	0.286
Observations:	5,916	5,916	5,916	5,916	5,916	5,916	5,916	5,916	5,916	5,916	5,916	5,916
Firm specifics:	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No

Note: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$ indicate the level of significance. Robust standard errors are used in the estimations. In the POLS column, the errors have been clustered by firm ID. The constant terms have been suppressed from the table output along with the year and industry dummy variables.

This sub-section provided the final part of the empirical analysis, which estimated the size of the productivity premium that multinationals enjoy compared to their domestic counterparts. The final chapter concludes all the findings and tested hypotheses.

6 Conclusion

This research paper analyses the productivity differences among Danish multinationals doing FDI across geographical regions and purely domestic firms, based on a sample of 2,119 industrialized firms, observed in the period 2009-2012. The purpose is to understand whether there are significant differences in TFP and efficiency growth between multinationals, depending on the host FDI regions. To my knowledge, this approach has not been applied previously in the empirical trade literature. The companies range from small to very large. The majority of the companies are purely domestic firms and selected few are companies that are engaged in foreign direct investments, which is a well seen empirical pattern in the literature.

The results provide evidence that there are significant differences between purely domestic firms and their counterparts with subsidiaries. As expected, MNEs have a much higher average productivity than non-MNEs. The same conclusion applies for the growth rates where firms with affiliates abroad have a productivity growth distribution to the right of domestic companies. The novel aspect of the paper is that it performs pairwise analysis of all the seven FDI regions in an attempt to uncover significant productivity differences. The Kolmogorov-Smirnov tests indicate that the productivity distributions of several pairs differ significantly. The cross-pair analysis also suggests that growth rates do not differ by region of FDI but levels of productivity do. Another important insight is the productivity ranking. It is shown that only the most productive firms are able to recover the costs to open subsidiaries in areas such as Africa and the Middle. Whereas the average productivity of firms having subsidiaries in Europe is lower, suggesting that the costs to open up an affiliate is lower so a bigger mass of firms can recuperate the fixed costs. On an aggregate level, without segmenting multinationals by host FDI region, both the growth and levels of TFP are significantly higher than these of purely domestic firms. This observation led to the interesting question whether growth converges among entrants and continuers on international markets. No support is found for the learning-by-doing hypothesis, because the TFP growth of entrants and continuers are not significantly different, suggesting that entering firms are not converging towards the existing firms doing FDI. As expected, the productivity levels and growth of entrants' and continuers' TFP are dominating the efficiency of purely domestic firms.

Finally, the investigation quantifies the FDI productivity premium over domestic companies. Relying on the Pooled OLS and Random effects estimates of the FDI premium model, on average multinationals enjoy a significant productivity advantage in the range of 15.4% - 19%, respectively. However, for only very few host regions additional premium is found. Across most specifications, the regional FDI dummies are insignificant. The only exceptions from the list are firms with with FDI in Europe, Asia and Africa for which both Pooled OLS and Random effects have positive significant values.

References

Alfaro, Laura, Areendam Chanda, Sebnem Kalemli-Ozcan, and Selin Sayek. 2004. "FDI and Economic Growth: The Role of Local Financial Markets." *Journal of International Economics*, 64(1): 89–112.

Arnold, Jens Matthias, and Katrin Hussinger. 2005. "Export Behavior and Firm Productivity in German Manufacturing: A Firm-level Analysis." *Review of World Economics*, 141(2): 219–243.

Aw, Bee Yan, Sukkyun Chung, and Mark J. Roberts. 2000. "Productivity and Turnover in the Export Market: Micro-level Evidence from the Republic of Korea and Taiwan (China)." *World Bank Economic Review*, 14(1): 65–90.

Aw, Bee Yan, Xiaomin Chen, and Mark J. Roberts. 2001. "Firm-level Evidence on Productivity Differentials and Turnover in Taiwanese Manufacturing." *Journal of Development Economics*, 66(1): 51–86.

Baldwin, John R., and Wulong Gu. 2003. "Export-Market Participation and Productivity Performance in Canadian Manufacturing." *Canadian Journal of Economics*, 36(3): 634–657.

Baltagi, Badi H., Peter Egger, and Michael Pfaffermayr. 2007. "Estimating Models of Complex FDI: Are There Third-Country Effects?" *Journal of Econometrics*, 140(1): 260–281.

Bernard, Andrew B., and J. Bradford Jensen. 1999. "Exceptional Exporter Performance: Cause, Effect, or Both?" *Journal of International Economics*, 47(1): 1–25.

Bernard, Andrew B., J. Bradford Jensen, and Robert Z. Lawrence. 1995. "Exporters, Jobs, and Wages in US Manufacturing: 1976-1987." *Brookings Papers on Economic Activity. Microeconomics*, 67–119.

Bernard, Andrew B., Jonathan Eaton, J. Bradford Jensen, and Samuel Kortum. 2003. "Plants and Productivity in International Trade." *The American Economic Review*, 93(4): pp. 1268–1290.

Blonigen, Bruce A. 2005. "A Review of the Empirical Literature on FDI Determinants." *Atlantic Economic Journal*, 33(4): 383–403.

Borensztein, Eduardo, Jose De Gregorio, and Jong-Wha Lee. 1998. "How Does Foreign Direct Investment Affect Economic Growth?" *Journal of International Economics*, 45(1): 115–135.

Brainard, S. Lael. 1997. "An Empirical Assessment of the Proximity-Concentration Trade-off between Multinational Sales and Trade." *The American Economic Review*, 87(4): 520–44.

Business, Doing. 2011. "Doing Business 2011: Making a difference for entrepreneurs." The World Bank Group.

Caves, Douglas W., Laurits R. Christensen, and Michael W. Tretheway. 1981. *U.S. Trunk Air Carriers, 1972-1977: A Multilateral Comparison of Total Factor Productivity.* Productivity Measurement in Regulated Industries.

Caves, Douglas W., Laurits R. Christensen, and W. Erwin Diewert. 1982. "The Economic Theory of Index Numbers and the Measurement of Input, Output, and Productivity." *Econometrica*, 50(6): 1393–1414.

Clerides, Sofronis K., Saul Lach, and James R. Tybout. 1998. "Is Learning By Exporting Important? Micro-Dynamic Evidence From Colombia, Mexico, And Morocco." *The Quarterly Journal of Economics*, 113(3): 903–947.

Conover, W. J. 1999. "Practical Nonparametric Statistics."

Damgaard, Jannick. 2011. "Productivity Spillovers from FDI: Ownership Structures, Domestic Firm Characteristics, and FDI Characteristics." Danmarks Nationalbank Working Papers.

Delgado, Miguel A., Jose C. Farinas, and Sonia Ruano. 2002. "Firms' Productivity and Export Markets: A Non-Parametric Approach." *Journal of International Economics*, 57(2): 397–422.

De Loecker, Jan. 2007. "Do Exports Generate Higher Productivity? Evidence from Slovenia." *Journal of International Economics*, 73(1): 69–98.

Eaton, Jonathan, Samuel Kortum, and Francis Kramarz. 2004. "Dissecting Trade: Firms, Industries, and Export Destinations." National Bureau of Economic Research, Inc.

Eberhardt, Markus, and Christian Helmers. 2010. *Untested Assumptions and Data Slicing: A Critical Review of Firm-level Production Function Estimators.* Department of Economics, University of Oxford.

Ekholm, Karolina, Rikard Forslid, and James R. Markusen. 2007. "Export-Platform Foreign Direct Investment." *Journal of the European Economic Association*, 776–795.

Faggio, Giulia, Kjell G. Salvanes, and John Van Reenen. 2010. "The Evolution of Inequality in Productivity and Wages: Panel Data Evidence." *Industrial and Corporate Change*, 19(6): 1919–1951.

Fons-Rosen, Christian, Sebnem Kalemli-Ozcan, Bent E. Sørensen, Carolina Villegas-Sanchez, and Vadym Volosovych. 2012. "Where are the Productivity Gains from Foreign Investment? Evidence on Spillovers and Reallocation from Firms, Industries and Countries." DEGIT, Dynamics, Economic Growth, and International Trade.

Fox, Jeremy T., and Valérie Smeets. 2011. "Does Input Quality Drive Measured Differences In Firm Productivity?" *International Economic Review*, 52(4): 961–989.

Girma, Sourafel, Holger Görg, and Eric Strobl. 2004. "Exports, International Investment, and Plant Performance: Evidence from a Non-Parametric Test." *Economics Letters*, 83(3): 317–324.

Girma, Sourafel, Richard Kneller, and Mauro Pisu. 2005. "Exports versus FDI: An Empirical Test." *Review of World Economics*, 141(2): 193–218.

Greenaway, David, and Richard Kneller. 2007. "Firm Heterogeneity, Exporting and Foreign Direct Investment." *The Economic Journal*, 117(517): F134–F161.

Griliches, Zvi, and Jacques Mairesse. 1995. "Production Functions: The Search for Identification." *NBER Working Paper*, , (5067).

Head, Keith, and John Ries. 2003. "Heterogeneity and the FDI versus Export Decision of Japanese Manufacturers." *Journal of the Japanese and International Economies*, 17(4): 448–467.

Helpman, Elhanan, Marc J. Melitz, and Stephen R. Yeaple. 2004. "Export versus FDI with Heterogeneous Firms." *The American Economic Review*, 300–316.

Hongye, Xiao, and Hao Feng. 2005. "The Applying of Capital Perpetual Inventory Method in China." *Finance & Trade Economics*, 3: 012.

Kimura, Fukunari, and Kozo Kiyota. 2006. "Exports, FDI, and Productivity: Dynamic Evidence from Japanese Firms." *Review of World Economics*, 142(4): 695–719.

Klette, Tor J., and Zvi Griliches. 1996. "The Inconsistency of Common Scale Estimators When Output Prices are Unobserved and Endogenous." *Journal of Applied Econometrics*, 11: 343–361.

Levinsohn, James, and Amil Petrin. 2003. "Estimating Production Functions Using Inputs to Control for Unobservables." *The Review of Economic Studies*, 70(2): 317–341.

Mairesse, Jacques, and Jordi Jaumandreu. 2005. “Panel-data Estimates of the Production Function and the Revenue Function: What Difference Does It Make?” *The Scandinavian Journal of Economics*, 107(4): 651–672.

Manova, Kalina, and Zhiwei Zhang. 2012. “Export Prices Across Firms and Destinations.” *The Quarterly Journal of Economics*, 127(1): 379–436.

Marschak, Jacob, and William H. Andrews. 1944. “Random Simultaneous Equations and the Theory of Production.” *Econometrica*, 143–205.

Meinen, Gerhard, Piet Verbiest, and P.P. de Wolf. 1998. “Perpetual Inventory Method.” *Service Lives, Discard Patterns, and Depreciation Methods. Heerlen-Voorburg: Statistics Netherlands.*

Melitz, Marc J. 2003. “The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity.” *Econometrica*, 71(6): 1695–1725.

Melitz, Marc J., and Stephen J. Redding. 2012. “Heterogeneous Firms and Trade.” National Bureau of Economic Research w18652.

Melitz, Marc J., Elhanan Helpman, and Stephen Yeaple. 2004. “Export Versus FDI with Heterogeneous Firms.” *The American Economic Review*, 94.

Navaretti, Giorgio Barba, and Davide Castellani. 2004. *Investments Abroad and Performance at Home: Evidence from Italian Multinationals.*

Nunnenkamp, Peter, and Julius Spatz. 2004. “FDI and Economic Growth in Developing Economies: How Relevant are Host-economy and Industry Characteristics?” *Transnational Corporations*, 13(3): 53–86.

Olley, G. Steven, and Ariel Pakes. 1996. “The Dynamics of Productivity in the Telecommunications Equipment Industry.” *Econometrica*, 64(6): 1263–1297.

Ottaviano, Gianmarco. 2011. “’New’ New Economic Geography: Firm Heterogeneity and Agglomeration Economies.” *Journal of Economic Geography*, 11(2): 231–240.

Petrin, Amil, Brian P. Poi, and James Levinsohn. 2004. “Production Function Estimation in Stata Using Inputs to Control for Unobservables.” *Stata Journal*, 4(2): 113–123.

Smeets, Valerie, and Frederic Warzynski. 2013. “Estimating Productivity with Multi-Product Firms, Pricing Heterogeneity and the Role of International Trade.” *Journal of International Economics*, 90(2): 237–244.

Temouri, Yama, Alexander Vogel, and Joachim Wagner. 2013. “Self-Selection into Export Markets by Business Services Firms–Evidence from France, Germany and the United Kingdom.” *Structural Change and Economic Dynamics*, 25: 146–158.

Tintelnot, Felix. 2013. “Global Production with Export Platforms.” Society for Economic Dynamics 2013 Meeting Papers 211.

Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data.* Vol. 1 of *MIT Press Books*, The MIT Press.

Appendix

A Profit maximizing behavior and price-setting of firms

The optimal price that a different type of firm has to charge is derived in order to maximize its profits. Consider from before that the demand for variety k, faced by a domestic producer is: $q(k) = Ap(k)^{-\epsilon}$; for exporters and multinationals it is: $q(k) = A^f p(k)^{-\epsilon}$, where A is the demand level in the home market, A^f is the demand level in the foreign country and $\epsilon = \frac{1}{(1-\alpha)}$. Furthermore, the incurred fixed costs by firms vary in their size. They are represented by f_D, f_X and f_{FDI} respectively and the following ranking holds: $f_D < f_X < f_{FDI}$. The variable production costs are dependent on productivity and are $\frac{c}{\theta(k)}$ for domestic manufacturers; $\frac{\tau c}{\theta(k)}$ for exporters and $\frac{c^f}{\theta(k)}$ per unit of output of variety k. Exporters incur iceberg trade costs and this is why their variable costs are higher by a factor $\tau > 1$.

A.1 Domestic firms

The optimal price for a firm that wants to sell a given variety k on the market is found by maximizing the profit equation for a domestic producer w.r.t. to $p(k)$. Following the introduced notation, the profit function of a firm willing to sell good k on the domestic market only is:

$$\pi_D(k) = p(k)q(k) - \frac{c}{\theta(k)}q(k) - f_D \tag{8}$$

Furthermore it is clear that firm profits depend on the productivity level, so I suppress the indexing by k where appropriate. Next, by substituting the expression for $q(k)$ into equation 8 one obtains:

$$\begin{aligned} \pi_D(\theta) &= p(k)Ap(k)^{-\epsilon} - \frac{c}{\theta}Ap(k)^{-\epsilon} - f_D \\ &= Ap(k)^{1-\epsilon} - \frac{Ac}{\theta}p(k)^{-\epsilon} - f_D \end{aligned}$$

The next step is to find the optimal price. This is done by taking the first derivative of the profit function w.r.t. to $p(k)$, which yields the following expression:

$$\begin{aligned} \frac{\partial \pi_D(\theta)}{\partial p(k)} &= \frac{\partial(Ap(k)^{1-\epsilon} - \frac{Ac}{\theta}p(k)^{-\epsilon} - f_j)}{\partial p(k)} = 0 \\ &= (1-\epsilon)Ap(k)^{-\epsilon} + \frac{Ac\epsilon}{\theta}p(k)^{-\epsilon-1} = 0 \end{aligned}$$

Moving the second expression to the other side of the equation and afterward diving by $p(k)^{-\epsilon}$ gives the following equation:

$$
\begin{aligned}
(1-\epsilon)Ap(k)^{-\epsilon} &= -\frac{Ac\epsilon}{\theta}p(k)^{-\epsilon-1} \\
(1-\epsilon)A &= -\frac{Ac\epsilon}{\theta}p(k)^{-1} \\
(1-\epsilon) &= -\frac{c\epsilon}{\theta p(k)} \\
p(k) &= -\frac{c\epsilon}{\theta(1-\epsilon)}
\end{aligned}
$$

Next, I substitute the expression for $\epsilon = \frac{1}{1-\alpha}$ from above to get:

$$
\begin{aligned}
p^{\star}(k) &= -\frac{c\left(\frac{1}{1-\alpha}\right)}{\theta\left(1-\left(\frac{1}{1-\alpha}\right)\right)} \\
&= -\frac{c\left(\frac{1}{1-\alpha}\right)}{\theta\left(\frac{-\alpha}{1-\alpha}\right)}
\end{aligned}
$$

$$
p^{\star}(k) = \frac{c}{\alpha\theta} \tag{9}
$$

which is the optimal price that firms willing to serve only the domestic market have to charge. The price charged for variety k is dependent on the variable production cost and the degree of substitution, which is assumed to be constant. ■

A.2 Exporters

In this subsection, the optimal price for a firm that wants to sell a given variety k on the export markets is derived. The profit function differs in its variable cost and fixed cost components. Exporters incur higher variable costs due to the trade expenditures, τ and also have higher fixed costs because for each new export market they incur additional costs such as market research and distribution expenditures. The profit function of a exporter is:

$$
\pi_X(k) = p(k)q(k) - \frac{\tau c}{\theta(k)}q(k) - f_X \tag{10}
$$

Now the expression for $q(k)$ is substituted into equation 10 one obtains:

$$
\begin{aligned}
\pi_X(\theta) &= p(k)A^f p(k)^{-\epsilon} - \frac{\tau c}{\theta}A^f p(k)^{-\epsilon} - f_X \\
&= A^f p(k)^{1-\epsilon} - \frac{A^f \tau c}{\theta}p(k)^{-\epsilon} - f_X
\end{aligned}
$$

The first derivative of the profit function w.r.t. to $p(k)$ is taken, which yields the following expression:

$$\begin{aligned}
\frac{\partial \pi_X(\theta)}{\partial p(k)} &= \frac{\partial (A^f p(k)^{1-\epsilon} - \frac{A^f \tau c}{\theta} p(k)^{-\epsilon} - f_X)}{\partial p(k)} = 0 \\
&= (1-\epsilon) A^f p(k)^{-\epsilon} + \frac{A^f \tau c \epsilon}{\theta} p(k)^{-\epsilon-1} = 0
\end{aligned}$$

Rearranging the expression and afterward diving by $p(k)^{-\epsilon}$ gives the following equation:

$$\begin{aligned}
(1-\epsilon) A^f p(k)^{-\epsilon} &= -\frac{A^f \tau c \epsilon}{\theta} p(k)^{-\epsilon-1} \\
(1-\epsilon) A^f &= -\frac{A^f \tau c \epsilon}{\theta} p(k)^{-1} \\
(1-\epsilon) &= -\frac{\tau c \epsilon}{\theta p(k)} \\
p(k) &= -\frac{\tau c \epsilon}{\theta (1-\epsilon)}
\end{aligned}$$

Finally, substituting $\epsilon = \frac{1}{1-\alpha}$ into the expression above gives us the following result:

$$\begin{aligned}
p^\star(k) &= -\frac{\tau c \left(\frac{1}{1-\alpha}\right)}{\theta \left(1 - \left(\frac{1}{1-\alpha}\right)\right)} \\
&= -\frac{\tau c \left(\frac{1}{1-\alpha}\right)}{\theta \left(\frac{-\alpha}{1-\alpha}\right)}
\end{aligned}$$

$$p^\star(k) = \frac{\tau c}{\alpha \theta} \tag{11}$$

which is the profit-maximing price that an exporter charges on the international market. Comparing equations 9 and 11 it can be seen that exporters charge a higher price due to the additional trade costs that they incur. ■

A.3 Multinationals

Finally, here I show the optimal price charged by a multinational firm that established production or subsidiary abroad. On one hand this type of firm incurs the highest fixed costs as they need to replicate their domestic investments abroad. On the other hand they save the variable trade costs, τ as they serve the foreign market directly. Thus, their profit function looks in the following way:

$$\pi_{FDI}(k) = p(k)q(k) - \frac{c^f}{\theta(k)} q(k) - f_{FDI} \tag{12}$$

where $\frac{c^f}{\theta}$ denotes the variable production cost per unit in the subsidiary. Firm profits depend on the productivity level, so I suppress the indexing by k where appropriate. Plugging in the expression for $q(k)$, faced by MNEs into equation 12 gives:

$$\begin{aligned}\pi_{FDI}(\theta) &= p(k)A^f p(k)^{-\epsilon} - \frac{c^f}{\theta}A^f p(k)^{-\epsilon} - f_{FDI} \\ &= A^f p(k)^{1-\epsilon} - \frac{A^f c^f}{\theta} p(k)^{-\epsilon} - f_{FDI}\end{aligned}$$

Again, I take the first derivative of the profit function, $\pi_{FDI}(\theta)$ w.r.t. to $p(k)$, which yields the following expression:

$$\begin{aligned}\frac{\partial \pi_{FDI}(\theta)}{\partial p(k)} &= \frac{\partial (A^f p(k)^{1-\epsilon} - \frac{A^f c^f}{\theta} p(k)^{-\epsilon} - f_{FDI})}{\partial p(k)} = 0 \\ &= (1-\epsilon)A^f p(k)^{-\epsilon} + \frac{A^f c^f \epsilon}{\theta} p(k)^{-\epsilon-1} = 0\end{aligned}$$

Rearranging the equation and diving by $p(k)^{-\epsilon}$ gives:

$$\begin{aligned}(1-\epsilon)A^f p(k)^{-\epsilon} &= -\frac{A^f c^f \epsilon}{\theta} p(k)^{-\epsilon-1} \\ (1-\epsilon)A^f &= -\frac{A^f c^f \epsilon}{\theta} p(k)^{-1} \\ (1-\epsilon) &= -\frac{c^f \epsilon}{\theta p(k)} \\ p(k) &= -\frac{c^f \epsilon}{\theta(1-\epsilon)}\end{aligned}$$

The last thing is to plug $\epsilon = \frac{1}{1-\alpha}$ in the last equation:

$$\begin{aligned}p^\star(k) &= -\frac{c^f \left(\frac{1}{1-\alpha}\right)}{\theta\left(1-\left(\frac{1}{1-\alpha}\right)\right)} \\ &= -\frac{c^f \left(\frac{1}{1-\alpha}\right)}{\theta\left(\frac{-\alpha}{1-\alpha}\right)}\end{aligned}$$

$$p^\star(k) = \frac{c^f}{\alpha\theta} \tag{13}$$

Respectively the profit-maximizing price for multinational firms that establish production facilities abroad is: $p^\star(k) = \frac{c^f}{\alpha\theta}$. ■

B Maximum profits of firms

From the Appendix A it became clear that the profit-maximizing prices domestic firms, exporters and multinationals charge are: $p(k) = \frac{c}{\alpha\theta}$, $p(k) = \frac{\tau c}{\alpha\theta}$ and $p(k) = \frac{c^f}{\alpha\theta}$ respectively. This section provides the maximized profits equation by plugging in the obtained optimal prices.

B.1 Domestic firms

First, by plugging the optimal price, $p(k) = \frac{c}{\alpha\theta}$, into the profit function $\pi_D(\theta) = p(k)q(k) - \frac{c}{\theta}q(k) - f_D = \left[p(k) - \frac{c}{\theta}\right] q(k) - f_D$:

$$\begin{aligned}
\pi_D^{\star}(\theta) &= \left[\underbrace{\frac{c}{\alpha\theta}}_{p^{\star}(k)} - \frac{c}{\theta}\right] \underbrace{A\left(\frac{c}{\alpha\theta}\right)^{-\epsilon}}_{q(k)} - f_D \\
&= \frac{c}{\theta}\left[\frac{1}{\alpha} - 1\right] Ac^{-\epsilon}\alpha^{\epsilon}\theta^{\epsilon} - f_D \\
&= c\theta^{-1}\left[\frac{1-\alpha}{\alpha}\right] Ac^{-\epsilon}\alpha^{\epsilon}\theta^{\epsilon} - f_D \\
&= \theta^{\epsilon-1}c^{1-\epsilon}A(1-\alpha)\alpha^{\epsilon-1} - f_D
\end{aligned}$$

$$= \theta^{\epsilon-1}\underbrace{(1-\alpha)A\left(\frac{c}{\alpha}\right)^{1-\epsilon}}_{D} - f_D = \theta^{\epsilon-1}D - f_D \tag{14}$$

where the profit function, which depends on the productivity level, is maximized for variety (firm) k and is represented in terms of the fixed costs f_D, an expression D and productivity θ. ■

B.2 Exporters

In the case of exporters the familiar profit function has the following form: $\pi_X(\theta) = p(k)q(k) - \frac{\tau c}{\theta}q(k) - f_D = \left[p(k) - \frac{\tau c}{\theta}\right] q(k) - f_X$. Next, the expression for exporter's optimal price is plugged into this profit function:

$$\begin{aligned}
\pi_X^{\star}(\theta) &= \left[\underbrace{\frac{\tau c}{\alpha\theta}}_{p^{\star}(k)} - \frac{\tau c}{\theta}\right] \underbrace{A^f\left(\frac{\tau c}{\alpha\theta}\right)^{-\epsilon}}_{q(k)} - f_X \\
&= \frac{\tau c}{\theta}\left[\frac{1}{\alpha} - 1\right] A^f\tau^{-\epsilon}c^{-\epsilon}\alpha^{\epsilon}\theta^{\epsilon} - f_X \\
&= \tau c\theta^{-1}\left[\frac{1-\alpha}{\alpha}\right] A^f\tau^{-\epsilon}c^{-\epsilon}\alpha^{\epsilon}\theta^{\epsilon} - f_X \\
&= \theta^{\epsilon-1}\tau^{1-\epsilon}c^{1-\epsilon}A^f(1-\alpha)\alpha^{\epsilon-1} - f_X
\end{aligned}$$

$$= \tau^{1-\epsilon}\theta^{\epsilon-1}\underbrace{(1-\alpha)A^f\left(\frac{c}{\alpha}\right)^{1-\epsilon}}_{D^f} - f_X = \tau^{1-\epsilon}\theta^{\epsilon-1}D^f - f_X \tag{15}$$

where the expression is the maximum profit expression, maximized for variety (firm) k and is represented in terms of the fixed costs f_X, an expression D^f and productivity θ. ■

B.3 Multinationals

The maximum profits equation for multinationals looks similar to the domestic producers' one but differs in the fixed costs and demand level, A^f. The optimal price of a multinational, $p^\star(k) = \frac{c^f}{\alpha\theta}$ is inserted into the profit function $\pi_{FDI}(\theta) = p(k)q(k) - \frac{c^f}{\theta}q(k) - f_{FDI} = \left[p(k) - \frac{c^f}{\theta}\right] q(k) - f_{FDI}$, where one obtains:

$$\begin{aligned}
\pi^\star_{FDI}(\theta) &= \left[\underbrace{\frac{c^f}{\alpha\theta}}_{p^\star(k)} - \frac{c^f}{\theta}\right] \underbrace{A^f \left(\frac{c^f}{\alpha\theta}\right)^{-\epsilon}}_{q(k)} - f_{FDI} \\
&= \frac{c^f}{\theta}\left[\frac{1}{\alpha} - 1\right] A{c^f}^{-\epsilon}\alpha^\epsilon\theta^\epsilon - f_{FDI} \\
&= c^f\theta^{-1}\left[\frac{1-\alpha}{\alpha}\right] A{c^f}^{-\epsilon}\alpha^\epsilon\theta^\epsilon - f_{FDI} \\
&= \theta^{\epsilon-1}{c^f}^{1-\epsilon}A^f(1-\alpha)\alpha^{\epsilon-1} - f_{FDI}
\end{aligned}$$

$$= \theta^{\epsilon-1}\underbrace{(1-\alpha)A^f\left(\frac{c^f}{\alpha}\right)^{1-\epsilon}}_{D^f_{FDI}} - f_{FDI} = \theta^{\epsilon-1}D^f_{FDI} - f_{FDI} \tag{16}$$

where the maximum profit function is represented in terms of the fixed costs f_{FDI}, an expression D^f_{FDI} and productivity θ.

C Data

This appendix focuses on the data characteristics and offers information on the firms in the dataset and the FDI flows and stocks by Danish industrialized firms. Summary statistics by internationalization status, size are also provided in order to assess the differences between purely domestic firms and multinationals.

In order to get a better sense of the size classification of firms, the ORBIS guidelines are provided here. Companies can be categorized in four distinct groups. The characteristics that have to be fulfilled are presented below.

C.1 Firm classification and variable construction

1. *Small:*

   (a) Total turnover $\leq 7,500,000$ DKK

   (b) Number of employees ≤ 15

   (c) Total assets $\leq 15,000,000$ DKK

   (d) Shall not belong to any of the other three groups

2. *Medium-sized:*

    (a) Total turnover $\geq 7,500,000$ DKK

    (b) Number of employees ≥ 15

    (c) Total assets $\geq 15,000,000$ DKK

    (d) Shall not be large or very large in size

3. *Large:*

    (a) Total turnover $\geq 75,000,000$ DKK

    (b) Number of employees ≥ 150

    (c) Total assets $\geq 150,000,000$ DKK

    (d) Shall not be very large in size

4. *Very large:*

    (a) Total turnover $\geq 750,000,000$ DKK

    (b) Number of employees ≥ 1000

    (c) Total assets $\geq 1,500,000,000$ DKK

    (d) Publicly listed

These rules have been applied in order to classify all the firms in the sample into four groups. Further insights about the variables in ORBIS, their measurement and source, can be found in the: official database user guide.. Next, the key variables in the production function are discussed.

Factors such as value-added, investments, material costs and capital are very important in the estimation of Cobb-Douglas production functions. The listed variables deserve a closer look as not all of them are directly observable in the dataset and thus had to be constructed from other entities.

All financial variables in the database, such as revenue (turnover), cost of goods sold, cost of employees, total assets, fixed assets, total tangible assets and capital are expressed in DKK, which are deflated to 2008 as base year in order to eliminate the effect of increasing prices, which could bias the true productivity estimates. In total three different types of deflators are applied on a NACE 2-digit industry level.[19]

1. Firm output:

    (a) Value-added: This measure is most often used as output in productivity estimations (Smeets and Warzynski, 2013; Griliches and Mairesse, 1995). It is defined as the gross revenue minus the material costs, $VA = TR - MTC$, where VA is value-added, TR is total revenue and MTC are material costs. This measure is not reported directly in ORBIS. The two variables that make up value-added are directly observed in the database. Evidence by Mairesse and

[19] The three different Producer Price Index (PPI) deflators for: Gross output, Value added and Intermediate inputs have been derived from the OECD STAN database (STAN08) for structural analysis. Unfortunately, the available years span up to 2010. In order to prolong the deflators up to 2012, it has been assumed that the PPI deflators increase yearly by the inflation rate. This rate for Denmark has been obtained from Statistics Denmark (PRIS9). The year 2008 is used as a base year for all PPI deflators.

Jaumandreu (2005) suggest that the impact of using firm-specific prices in the process of deflating company revenues, has little impact on the obtained coefficients. Nevertheless, a value added PPI deflator is applied to convert nominal values in real terms.

(b) Revenue (turnover): The revenue measure is directly observed and captures the annual total sales in thousands of DKK of a given firm. Since ORBIS does not contain information on the output prices of firms, the nominal values have been deflated using a gross output PPI. The measure can be used as a dependent variable in the production function estimation. Using revenues as gross output underlines that firms operate in a perfectly competitive market. If firms produce differentiated products then a significant dispersion in output prices can be seen, which is moreover correlated with the observable inputs. This is why deflated revenues may not be the most suitable measure for output and would lead to a downward bias in the estimates (Klette and Griliches, 1996).

2. Firm intermediate inputs:

   (a) Labor: The labor input is represented by the number of employees working in a given company on an annual basis. The variable is directly reported by ORBIS. Unfortunately, using the number of employees does not say anything about the quality of the labor input. Some of the reported employees could be blue-collar workers whereas others could be white-collar workers. Different levels of education could also affect the productivity of employees. An interesting suggestion has been offered by Fox and Smeets (2011), where the cost of employees is used as a proxy to control for the quality of the input factor. In that way the labor input can be measured as the total wage bill. This measure is available in the dataset and is used as a robustness check for the obtained results. Another potential improvement to the labor measure could be to use the total number of worked hours, as also expressed by Eberhardt and Helmers (2010). It would be an even more accurate measure of how much labor actually goes into one unit of output. Such a detailed variable is not available so that the number of employees has been used as a second-best option.

   (b) Capital: Even though this variable exists directly in the database, it should be noted that it represents equity capital (Nominal capital + Capital of general partners + Capital of limited partners) of a given company. The measure that is needed in the empirical analyses is productive assets - those that provide services in the output generating process. For that reason physical capital is proxied by using the reported variable for "total assets", which captures entities such as buildings, plants, cars, land and other tangible assets. The variable is directly reported on an annual basis and is in nominal terms. It is transformed in real values by using the gross output PPI deflator on a 2-digit NACE industry level.

3. Proxy variables in the production function:

   (a) Cost of goods sold: This is a collective variable, which is directly reported in ORBIS. The values are available in DKK and are deflated using an industry-specific PPI for intermediate inputs. The variable is used as an equivalent measure to intermediate inputs that can proxy for the unobserved productivity shocks in the method, suggested by Levinsohn and Petrin (2003). The idea is to capture all the costs of intermediate inputs involved in the production, such as materials, energy and packaging. Since the database does not provide separate information on each of these variables, it is decided to rely upon the cost of goods sold as an aggregate measure for the inputs entering production.

(b) Investments: In the original database, values for investments are very rare and for the Danish sub-sample they are almost entirely missing. The variable is very important for the Olley and Pakes (1996) productivity estimation method, which is introduced in Section 3.1. Thus, obtaining a good proxy for investments is crucial. To resolve this issue, the measure is derived by using the perpetual inventory method (PI).[20] which relies strongly on capital availability. The assumption is that investments should necessarily be equal to the difference between the capital today and last year's non-depreciated capital: $I_t = K_{t+1} - (1 + \delta)K_t$, where I_t and K_t are the investment and capital levels at time t, respectively and δ is the average yearly depreciation rate applied to capital. An average rate of 10% is used, even though different asset classes have varying depreciation levels.[21] In the process, one year of observations are because the investment dynamics involve two years. The PI approach offers a partial solution to the missing data for investments (less than 4% of the firms have reported investment values), but it has to be emphasized that the lack of information about this variable is a serious caveat of the ORBIS database.

4. Additional variables:

   (a) Number and location of subsidiaries: These variables play a very important role in analysis of the FDI-productivity link. For each company, the number of subsidiaries are available on a regional and yearly basis. Furthermore, subsidiary dummy variables are created, which indicate whether a company has any foreign affiliates or not in a specific geographical region. Subsidiaries in Denmark are excluded from the count.

   (b) Industry and size: These categorical variables are used as controls in the TFP estimations and segment firms into different industries and sizes.

All the monetary variables are also log-transformed as this is the required format for the estimation of the production function. This allows the obtained coefficients to be interpreted as elasticities.

C.2 Distribution of firms

The distribution of firms by size and year before the data cleaning process is shown in Table 16. On one hand, "Before data cleaning" includes only firm observations that have non-missing revenue (turnover) values in the dataset. This results in an unbalanced panel since in some years revenues are not reported, especially for some small firms. Applying this data cleaning criteria, leaves 24,709 observations for 8,296 companies from the starting number of 129,921 observations for 26,285 firms. This cleaning procedure is absolutely necessary since revenue data is missing for 80.4% of the sample. The turnover variable plays a focal role in the econometric exercises and productivity estimation so it

[20] For more information on the method, see Meinen, Verbiest and Wolf (1998). For an overview of the difficulties arising when estimating investment and capital levels, see Hongye and Feng (2005). Some studies avoid using the method by using the argument that capital levels may suffer from measurement error, which would then bias the correct calculation of the investment variable.

[21] The approximate average rate is taken from the taxation and investment report by Deloitte Denmark. For reference purposes, 9.3% is the depreciation rate of capital in the UK, suggested by the UK Office for National Statistics.

is a requirement that at least revenue streams are available for all observations. On the other hand, “After data cleaning” is the version of the dataset that includes only firm observations that have non-missing values for revenue (turnover), cost of goods sold, number of employees, Industry classification, capital, total tangible assets and number of subsidiaries. The latter criteria imposes more requirements on the selection, which naturally leads to fewer observations. Table 2 in the main body of the text provides the distribution by size and year after the data has been cleaned.

A further difference is that Table 16 provides descriptive statistics for an unbalanced version of the panel, whereas Table 16 displays the distribution of firms by size and year in a balanced panel. In the former case, some firms have missing values on certain variables in some years, meaning that there are time gaps within each panel. In the latter situation only firm observations that contain non-missing values for the specified variables across all four years are kept.

Comparing the percentages in the “before data cleaning” columns of both tables, reveals that after the more demanding segmentation criteria was introduced, the share of medium, large and very large firms increased in comparison to small companies. In the balanced panel, summarized in Table 16, small firms now represent 30.10%, medium-sized occupy 21.54%, large ones have 30.83% and very large companies take the remaining 17.53% of the sample.

Table 16: Firm segmentation by size and year before data cleaning (2009-2012)

Dataset version	Balanced panel				Unbalanced panel			
Size \ Year	2009	2010	2011	2012	2009	2010	2011	2012
Small	30.10%	30.10%	30.10%	30.10%	42.37%	40.76%	39.27%	40.54%
Medium-sized	21.54%	21.54%	21.54%	21.54%	20.93%	20.60%	20.31%	19.12%
Large	30.83%	30.83%	30.83%	30.83%	23.74%	24.94%	26.13%	26.30%
Very large	17.53%	17.53%	17.53%	17.53%	12.96%	13.70%	14.28%	14.03%
Total:	100%	100%	100%	100%	100%	100%	100%	100%
Observations:	4,359	4,359	4,359	4,359	6,103	6,175	6,203	6,228

Note: The statistics are based on a dataset before cleaning, corresponding to stage 2.

After cleaning the data further by keeping only observations with non-missing values for a group of variables, it is noticeable that the number of observations declines. In Table 16, the distribution by size for 2012 seems to differ from the general trend under the column “after data cleaning”. The number of observations are at their lowest and in the meanwhile the shares of very large and large companies increase substantially. After balancing the panel and only keeping firms with full data coverage for the period 2009-2012, the company distribution is as follows: small establishments (18.59%), medium-sized firms (20.58%), large companies (39.36%) and finally very large organizations (21.47%). The total number of observations in the balanced panel after the data cleaning is 8,476 (Table 3).

Table 17 shows the key variable names in the database and their description. The dataset variables available on a disc upon request.

Table 17: Names and definitions of database variables

Dataset var.	Definition	Dataset var.	Definition
$NACE\_ind$	2-digit Industry	$dl\_Revenue$	Deflated revenues
$Nemp$	Number of employees	$dl\_Cogsold$	Deflated cost of goods sold
$Categ$	Size category	dl_Vadded	Deflated value-added
$Nsubs$	Number of subsidiaries abroad (total)	$dl\_Tassets$	Deflated total tangible assets
$nsub\_EU$	Number of subsidiaries in Europe	dl_Inv	Deflated investments
$nsub\_ME$	Number of subsidiaries in the Middle East	$dl\_Fassets$	Deflated fixed assets
$nsub\_ASP$	Number of subsidiaries in Asia Pacific	$dl\_Capital$	Deflated capital (equity)
$nsub\_AS$	Number of subsidiaries in Asia	dl_Cemp	Deflated cost of employees
$nsub\_NA$	Number of subsidiaries in North America	s_emp	Labor cost share of revenues
$nsub\_SA$	Number of subsidiaries in South America	s_cap	Capital cost share of revenues
$nsub\_AF$	Number of subsidiaries in Africa	$status$	International status: MNE or purely domestic

C.3 FDI intensity and export development

Table 18 presents yearly data for the FDI undertaken by Danish firms, the accumulated stock of direct investments abroad and also export development from Denmark to the rest of the world by regions. The information in the table is visualized in Figures 1, 2.

Table 18: Outward stock and flows of FDI and exports in bil. DKK (2004-2012)

		2004	2005	2006	2007	2008	2009	2010	2011	2012
FDI stock from DK	to World	579.3	677	723.3	798.3	899.7	973.7	1145	1221.5	1304.3
	to EU	348.1	427.9	447.9	451.5	554	581.2	665.7	706.7	775.3
	to Africa	5.7	6.0	6.8	6.0	6.0	7.2	10.8	12.8	16.4
	to America	73.7	81.9	89.7	94.4	98.7	121.2	156.3	166.4	167.3
	to Asia	22.3	26.0	31.7	42.1	54.5	54.6	63.9	66.8	67.0
	to Oceania	3.0	4.0	6.5	7.8	9.4	13.1	15.2	15.2	18.1
FDI flows from DK	to World	-66.9	97.1	48.5	112.2	67.6	33.9	-0.8	67.7	46.2
	to Europe	-71.8	89.4	22.5	63.4	60.0	10.0	-15.9	50.6	43.6
	to Africa	1.8	-3.5	13.1	32.0	0.4	14.6	6.5	1.4	-2.0
	to America	0.1	0.2	2.5	0.7	-3.8	0.6	-8.8	-0.8	5.9
	to Asia	2.0	7.6	10.2	17.7	11.2	7.2	15.2	15.3	-1.2
	to Oceania	1.1	2.1	0.0	-1.7	-0.4	1.3	2.6	1.2	-0.1
Export from DK	to World	452.4	490.7	538.3	551.6	586.9	490.1	538.3	600.1	614.7
	to EU	314.6	339.4	373.4	379.9	402.4	323.1	346.5	385.7	383.1
	to Africa	4.0	4.2	4.9	5.0	6.3	4.8	5.7	6.5	7.9
	to America	32.3	38.3	44.7	45.2	46.3	42.3	48.5	51.2	52.7
	to Asia	42.1	44.2	47.1	49.0	57.0	50.0	59.0	71.7	72.6
	to Oceania	4.6	4.0	4.6	4.9	4.7	5.2	5.6	8.1	7.3

Sources: Statistics Denmark, 2014, Variables: DNDIRA2 and KN8Y; OECD Statistics, 2014, Variable: FDI flows by partner country

D Additional estimation results

This appendix offers the key results of the productivity estimations and empirical FDI exercises.

Table 19: Levinsohn-Petrin (LP_TFP) estimates for Industries 1-6 (2009-2012)

Log of value-added, (VA)						
Industry	Agriculture	Mining	Manufacturing	Electricity, gas	Water supply	Construction
Classification	Ind. 1	Ind. 2	Ind. 3	Ind. 4	Ind. 5	Ind. 6
	β	β	β	β	β	β
Log of Labor, (L)	0.589*	0.256	0.577***	-0.150	0.089	0.466***
	(0.32)	(0.30)	(0.05)	(0.14)	(0.22)	(0.09)
Log of Capital (Assets), (K)	0.477**	0.723**	0.646***	0.415	0.469	0.639***
	(0.21)	(0.34)	(0.09)	(0.27)	(0.30)	(0.12)
$L + K$	1.066	0.979	1.223	0.265	0.558	1.105
Wald test for CRS, χ^2 statistic	0.02	0.00	3.82*	3.86**	1.70	0.75
Constant returns to scale	Yes	Yes	No	No	Yes	Yes
Observations	46	56	1,628	89	63	632

Note: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$ indicate the level of significance. The standard errors of the estimates are reported in parentheses under the β coefficients.

Table 20: Levinsohn-Petrin (LP_TFP) estimates for Industries 7-13 (2009-2012)

Log of value-added, (VA)						
Industry	Retail trade	Transportation	Accommodation	Info tech.	Finance	Pro. activities
Classification	Ind. 7	Ind. 8	Ind. 9	Ind. 10	Ind. 11	Ind. 13
	β	β	β	β	β	β
Log of Labor, (L)	0.462***	0.496***	1.125***	0.641***	0.498***	0.779***
	(0.05)	(0.08)	(0.26)	(0.06)	(0.06)	(0.08)
Log of Capital (Assets), (K)	0.415***	1.007***	-0.092	0.339***	0.840***	0.621***
	(0.06)	(0.32)	(0.34)	(0.12)	(0.18)	(0.14)
$L + K$	0.877	1.503	1.033	0.98	1.338	1.4
Wald test for CRS, χ^2 statistic	3.92**	2.43	0.01	0.02	1.60	8.41***
Constant returns to scale	No	Yes	Yes	Yes	Yes	No
Observations	2,286	346	122	639	541	864

Note: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$ indicate the level of significance. The standard errors of the estimates are reported in parentheses under the β coefficients.

Table 21: Levinsohn-Petrin (LP_TFP) estimates for Industries 14-19 (2009-2012)

Log of value-added, (VA)					
Industry	Admin. services	Education	Human health	Arts	Other services
Classification	Ind. 14	Ind. 16	Ind. 17	Ind. 18	Ind. 19
	β	β	β	β	β
Log of Labor, (L)	0.446***	1.486**	0.610***	0.509***	0.694
	(0.06)	(0.65)	(0.16)	(0.14)	(0.49)
Log of Capital (Assets), (K)	0.595***	0.074	0.348	0.344**	0.493
	(0.13)	(0.38)	(0.40)	(0.15)	(0.40)
$L+K$	1.041	1.56	0.958	0.853	1.187
Wald test for CRS, χ^2 statistic	0.07	0.42	0.01	0.54	0.21
Constant returns to scale	Yes	Yes	Yes	Yes	Yes
Observations	297	16	144	68	60

Note: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$ indicate the level of significance. The standard errors of the estimates are reported in parentheses under the β coefficients.

The set of tables contain the estimation of a Cobb-Douglas production function with two inputs by using FE estimation.

Table 22: Fixed effects (OLS_TFP) estimates for Industries 1-6 (2009-2012)

Log of value-added, (VA)						
Industry	Agriculture	Mining	Manufacturing	Electricity, gas	Water supply	Construction
Classification	Ind. 1	Ind. 2	Ind. 3	Ind. 4	Ind. 5	Ind. 6
	β	β	β	β	β	β
Log of Labor, (L)	0.419	0.343*	0.553***	0.356**	-0.096	0.468***
	(0.35)	(0.18)	(0.11)	(0.17)	(0.33)	(0.17)
Log of Capital (Assets), (K)	0.298	0.346	0.329***	0.255	0.543***	0.343***
	(0.36)	(0.20)	(0.08)	(0.24)	(0.11)	(0.12)
$L+K$	0.717	0.689	0.882	0.611	0.447	0.811
Wald test for CRS, F statistic	0.32	1.91	2.41	2.28	2.49	1.63
Constant returns to scale	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	46	56	1,628	89	63	632

Note: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$ indicate the level of significance. The standard errors of the estimates are reported in parentheses under the β coefficients. The constant term has been suppressed from the table output along with the Year FE.

Table 23: Fixed effects (OLS_TFP) estimates for Industries 7-13 (2009-2012)

Log of value-added, (VA)						
Industry	Retail trade	Transportation	Accommodation	Info tech.	Finance	Pro. activities
Classification	Ind. 7	Ind. 8	Ind. 9	Ind. 10	Ind. 11	Ind. 13
	β	β	β	β	β	β
Log of Labor, (L)	0.318***	0.388**	0.659***	0.455***	0.738***	0.395***
	(0.08)	(0.18)	(0.13)	(0.12)	(0.10)	(0.09)
Log of Capital (Assets), (K)	0.300***	0.467***	0.343**	0.229	0.146	0.526***
	(0.04)	(0.13)	(0.15)	(0.17)	(0.11)	(0.10)
$L + K$	0.618	0.855	1.002	0.684	0.884	0.921
Wald test for CRS, F statistic	20.10***	0.81	0.00	3.50*	1.09	0.54
Constant returns to scale	No	Yes	Yes	No	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,304	347	124	644	541	881

Note: *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$ indicate the level of significance. The standard errors of the estimates are reported in parentheses under the β coefficients. The constant term has been suppressed from the table output along with the Year FE.

Table 24: Fixed effects (OLS_TFP) estimates for Industries 14-19 (2009-2012)

Log of value-added, (VA)					
Industry	Admin. services	Education	Human health	Arts	Other services
Classification	Ind. 14	Ind. 16	Ind. 17	Ind. 18	Ind. 19
	β	β	β	β	β
Log of Labor, (L)	0.146	1.539	0.134*	0.659**	0.063
	(0.09)	(0.76)	(0.07)	(0.25)	(0.16)
Log of Capital (Assets), (K)	0.357**	0.142	-0.009	0.201	0.162
	(0.16)	(0.12)	(0.17)	(0.12)	(0.17)
$L + K$	0.503	1.681	0.125	0.86	0.225
Wald test for CRS, F statistic	11.59***	0.73	23.41***	0.39	146.52***
Constant returns to scale	No	Yes	No	Yes	No
Year FE	Yes	Yes	Yes	Yes	Yes
Observations	304	16	145	68	60

Note: *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$ indicate the level of significance. The standard errors of the estimates are reported in parentheses under the β coefficients. The constant term has been suppressed from the table output along with the Year FE.

The CRS hypothesis is one-sided, meaning that it does not tell us whether there are increasing or decreasing returns to scale. It simply tests whether there are *constant returns to scale.* Thus, in some occasions the coefficients are very far away from unity, meaning that most likely the industry is experiencing decreasing returns to scale and vice versa. The statistical tests shall be understood with caution. Some industries have very small number of observations, which makes it hard to draw conclusive inference about the significance of the returns.

Table 25 below presents the correlation matrix among the different productivity measures to assess the magnitude of results. Overall a positive correlation is observed among all the measures, which is a positive sign. Thus, results should have comparable magnitudes and similar signs.

Table 25: Correlation matrix of different TFP measures for the six main industries (2009-2012)

	Index TFP	Levinsohn and Petrin TFP	Labor TFP	Olley and Pakes TFP
	I_TFP	LP_TFP	L_TFP	OP_TFP
I_TFP	1.000			
LP_TFP	0.369	1.000		
L_TFP	0.218	0.420	1.000	
OP_TFP	0.429	0.504	0.624	1.000

D.1 Productivity evolution

The impact of chain-linking in the multilateral TFP index is visible in Figures 13 and 14.

Figure 13: Chain-linked vs. unlinked Index TFP of MNEs, 2012

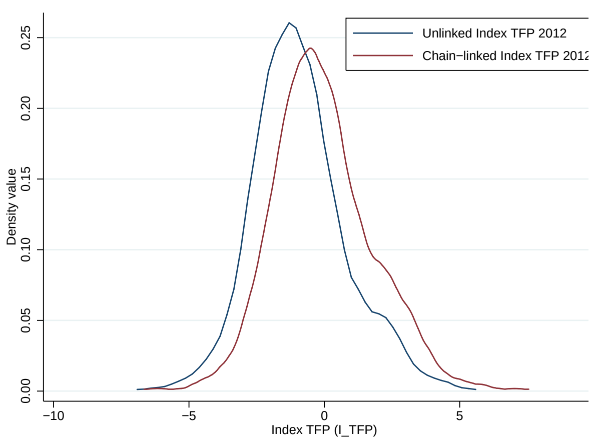
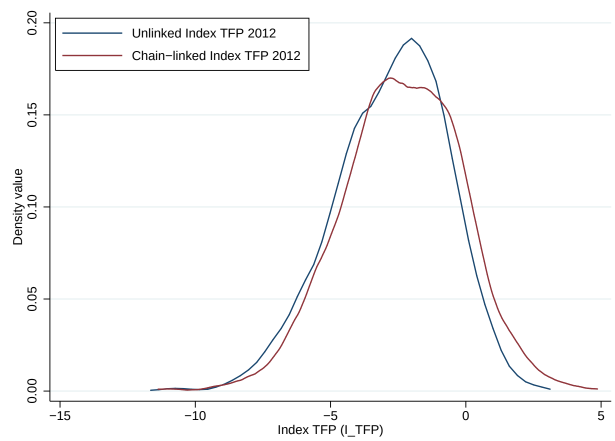


Figure 14: Chain-linked vs. unlinked Index TFP of purely domestic companies, 2012



The following set of matrix tables contain the detailed K-S tests performed on the distribution of Index TFP by regional pairs. The productivity distributions are benchmarked against each other. This analysis is performed both for the Index TFP but also for the change (growth) in the TFP across the period of interest. Tables 26 and 27 include the estimates for the differences between TFP distributions, whereas Tables 28 and 29 analyze the distribution of efficiency change on a cross-pair basis. The tables are understood as follows:

- The values under the gray diagonal are the results from the K-S test for 2009 (Table 26) and 2011 (Table 27), respectively. The values above the diagonal in both tables report the results of the K-S tests for 2010 (Table 26) and 2012 (Table 27), respectively.
- “K-S stat.” shows the Kolmgorov-Smirnov test statistic for the given pair.
- The cell “Diff.” states whether one of the TFP distributions stochastically dominates the other one in the pair at a standard significance level of 5%.

For example, looking at Table 26, it can be stated that there are significant differences in the distribution of efficiency levels between firms that do FDI in Europe and those that do in the Middle East.

Table 26: Kolmogorov-Smirnov tests for TFP distributions of Non-MNEs vs. MNEs with subsidiaries by region for 2009 and 2010

			2010													
			EU		ME		ASP		AS		NA		SA		AF	
			K-S stat.	Diff.	K-S stat.	Diff.	K-S stat.	Diff.	K-S stat.	Diff.	K-S stat.	Diff.	K-S stat.	Diff.	K-S stat.	Diff.
2009	EU	K-S stat.			0.430***	Yes	0.201***	Yes	0.301***	Yes	0.170***	Yes	0.309***	Yes	0.499***	Yes
		Diff.														
	ME	K-S stat.	0.360***				0.306	No	0.188	No	0.228	No	0.202	No	0.366**	Yes
		Diff.	Yes													
	ASP	K-S stat.	0.238***		0.260				0.267***	Yes	0.130	No	0.232**	Yes	0.428***	Yes
		Diff.	Yes		No											
	AS	K-S stat.	0.438***		0.293*		0.353***				0.193	No	0.145	No	0.301**	Yes
		Diff.	Yes		No		Yes									
	NA	K-S stat.	0.142**		0.182		0.134		0.157				0.255***	Yes	0.456***	Yes
		Diff.	Yes		No		No		No							
	SA	K-S stat.	0.313***		0.150		0.220**		0.144		0.298***				0.409***	Yes
		Diff.	Yes		No		Yes		No		Yes					
	AF	K-S stat.	0.493***		0.431***		0.410***		0.234		0.476***		0.400***			
		Diff.	Yes		Yes		Yes		No		Yes		Yes			

Note: *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$ indicate the level of significance of the K-S statistic, which is the maximum difference between the two tested distributions. The differences in the distribution of TFP levels (Yes/No) are evaluated at a 95% confidence level, meaning that if the p-value exceeds 0.05, then it is considered that there are is *no* difference in the productivity distributions in a given pair.

The test statistic is 0.360 and is significant at a 1% level. Alternatively, for 2010, it can be stated that the TFP distribution of multinationals with subsidiaries in Asia Pacific is *not equal* to the efficiency distribution of companies doing FDI in Europe. The test statistic is 0.201 and is again significant at a 1% level. The two examples showcase how Table 26 through 29 should be interpreted.

FDI premium estimation using alternative TFP measure. The results below are for the Labor productivity construct, which is reported as a robustness measure.

Table 27: Kolmogorov-Smirnov tests for TFP distributions of Non-MNEs vs. MNEs with subsidiaries by region for 2011 and 2012

			2012													
			EU		ME		ASP		AS		NA		SA		AF	
			K-S stat.	Diff.	K-S stat.	Diff.	K-S stat.	Diff.	K-S stat.	Diff.	K-S stat.	Diff.	K-S stat.	Diff.	K-S stat.	Diff.
2011	EU	K-S stat.			0.412***	Yes	0.150***	Yes	0.308***	Yes	0.137**	Yes	0.233***	Yes	0.428***	Yes
		Diff.														
	ME	K-S stat.	0.389***				0.266	No	0.127	No	0.345**	Yes	0.201	No	0.293*	No
		Diff.	Yes													
	ASP	K-S stat.	0.194***		0.361**				0.333***	Yes	0.160	No	0.222**	Yes	0.408***	Yes
		Diff.	Yes		Yes											
	AS	K-S stat.	0.342***		0.172		0.309***				0.158	No	0.107	No	0.276**	Yes
		Diff.	Yes		No		Yes									
	NA	K-S stat.	0.148***		0.322*		0.083		0.171				0.228**	Yes	0.433***	Yes
		Diff.	Yes		No		No		No							
	SA	K-S stat.	0.289***		0.223		0.233**		0.160		0.257***				0.404***	Yes
		Diff.	Yes		No		Yes		No		Yes					
	AF	K-S stat.	0.432***		0.354**		0.407***		0.262**		0.448		0.392***			
		Diff.	Yes		Yes		Yes		Yes		Yes		Yes			

Note: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$ indicate the level of significance of the K-S statistic, which is the maximum difference between the two tested distributions. The differences in the distribution of TFP levels (Yes/No) are evaluated at a 95% confidence level, meaning that if the p-value exceeds 0.05, then it is considered that there are is *no* difference in the productivity distributions in a given pair.

Table 28: Kolmogorov-Smirnov tests for the TFP growth rate distributions of Non-MNEs vs. MNEs with subsidiaries by region for 2010 and 2011

			2011													
			EU		ME		ASP		AS		NA		SA		AF	
			K-S stat.	Diff.	K-S stat.	Diff.	K-S stat.	Diff.	K-S stat.	Diff.	K-S stat.	Diff.	K-S stat.	Diff.	K-S stat.	Diff.
2010	EU	K-S stat.			0.124	No	0.099	No	0.076	No	0.085	No	0.100	No	0.191*	No
		Diff.														
	ME	K-S stat.	0.122				0.194	No	0.163	No	0.184	No	0.189	No	0.243	No
		Diff.	No													
	ASP	K-S stat.	0.069		0.248				0.109	No	0.216**	Yes	0.138	No	0.151	No
		Diff.	No		No											
	AS	K-S stat.	0.062		0.129		0.090				0.177	No	0.147	No	0.173	No
		Diff.	No		No		No									
	NA	K-S stat.	0.072		0.208		0.143		0.144				0.134	No	0.219*	No
		Diff.	No		No		No		No							
	SA	K-S stat.	0.165*		0.244		0.209**		0.173		0.157				0.170	No
		Diff.	No		No		Yes		No		No					
	AF	K-S stat.	0.079		0.165		0.064		0.081		0.097		0.194			
		Diff.	No		No		No		No		No		No			

Note: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$ indicate the level of significance of the K-S statistic, which is the maximum difference between the two tested distributions. The differences in the distribution of TFP growth rates (Yes/No) are evaluated at a 95% confidence level, meaning that if the p-value exceeds 0.05, then it is considered that there are is *no* difference in distribution of TFP growth in a given pair.

Table 29: Kolmogorov-Smirnov tests for the TFP growth rate distributions of Non-MNEs vs. MNEs with subsidiaries by region for 2012

			EU		ME		ASP		AS		NA		SA		AF	
			K-S stat.	Diff.	K-S stat.	Diff.	K-S stat.	Diff.	K-S stat.	Diff.	K-S stat.	Diff.	K-S stat.	Diff.	K-S stat.	Diff.
2012	EU	K-S stat.														
		Diff.														
	ME	K-S stat.	0.249***													
		Diff.	Yes													
	ASP	K-S stat.	0.123**		0.222											
		Diff.	Yes		No											
	AS	K-S stat.	0.110		0.193		0.076									
		Diff.	No		No		No									
	NA	K-S stat.	0.134**		0.122		0.210**		0.176							
		Diff.	Yes		No		Yes		No							
	SA	K-S stat.	0.107		0.176		0.069		0.075		0.099					
		Diff.	No		No		No		No		No					
	AF	K-S stat.	0.125		0.108		0.096		0.081		0.074		0.128			
		Diff.	No		No		No		No		No		No			

Note: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$ indicate the level of significance of the K-S statistic, which is the maximum difference between the two tested distributions. The differences in the distribution of TFP growth rates (Yes/No) are evaluated at a 95% confidence level, meaning that if the p-value exceeds 0.05, then it is considered that there are is *no* difference in distribution of TFP growth in a given pair.

Table 30: FDI premiums - POLS, FE and RE estimates of Labor TFP levels

θ	Pooled OLS				Fixed Effects				Random Effects			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
FDI premium	0.171***	0.115***			0.012	-0.003			0.134***	0.059*		
	(0.03)	(0.02)			(0.04)	(0.05)			(0.04)	(0.03)		
Age	0.003***		0.002***		-0.001		-0.001		0.005***		0.005***	
	(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)	
Log employees	-0.039***		-0.041***		-0.479***		-0.480***		-0.099***		-0.100***	
	(0.01)		(0.01)		(0.05)		(0.05)		(0.02)		(0.02)	
MNE (EU)			0.089***	0.039			-0.065	-0.079			0.042	-0.02
			(0.03)	(0.02)			(0.05)	(0.05)			(0.04)	(0.04)
MNE (ME)			-0.009	-0.031			0.044	0.01			0.035	0.009
			(0.08)	(0.08)			(0.05)	(0.05)			(0.04)	(0.04)
MNE (ASP)			0.031	0.006			-0.013	0.003			0.034	0.013
			(0.06)	(0.06)			(0.04)	(0.05)			(0.04)	(0.04)
MNE (AS)			0.096*	0.08			0.077*	0.055			0.094**	0.073*
			(0.06)	(0.06)			(0.04)	(0.04)			(0.04)	(0.04)
MNE (NA)			0.011	0.006			0.044	0.025			0.065*	0.036
			(0.04)	(0.04)			(0.04)	(0.04)			(0.04)	(0.04)
MNE (SA)			0.240***	0.213 **			0.045	0.052			0.139**	0.106*
			(0.08)	(0.09)			(0.05)	(0.07)			(0.06)	(0.06)
MNE (AF)			0.135	0.145			0.008	0.024			0.065	0.059
			(0.09)	(0.09)			(0.05)	(0.06)			(0.06)	(0.06)
R^2	0.121	0.115	0.123	0.117	0.068	0.004	0.069	0.005	0.108	0.114	0.109	0.115
Observations:	7,885	7,885	7,885	7,885	7,885	7,885	7,885	7,885	7,885	7,885	7,885	7,885
Firm specifics:	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No

Note: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$ indicate the level of significance. Robust standard errors are used in the estimations. In the POLS column, the errors have been clustered by firm ID. The constant terms have been suppressed from the table output along with the year and industry dummy variables.

Table 31: FDI premiums - POLS, FE and RE estimates of Labor TFP change

$\Delta\theta$	Pooled OLS				Fixed Effects				Random Effects			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
FDI premium	0.009	0.009			0.086	0.068			0.009	0.009		
	(0.02)	(0.01)			-0.06	(0.07)			(0.01)	(0.01)		
Age	0.001		0.001		-0.045**		-0.044**		-0.000		-0.000	
	(0.00)		(0.00)		(0.01)		(0.01)		(0.00)		(0.00)	
Log employees	0.001		0.001		-0.537**		-0.538**		0.000		0.001	
	(0.00)		(0.00)		(0.10)		(0.10)		(0.00)		(0.00)	
MNE (EU)			0.011	0.012			0.044	0.030			0.011	0.012
			(0.02)	(0.02)			(0.08)	(0.09)			(0.01)	(0.01)
MNE (ME)			0.007	0.008			0.047	0.007			0.007	0.008
			(0.03)	(0.03)			(0.07)	(0.06)			(0.03)	(0.03)
MNE (ASP)			-0.072**	-0.072**			-0.165	-0.150			-0.072***	-0.072***
			(0.04)	(0.03)			(0.14)	(0.14)			(0.03)	(0.03)
MNE (AS)			0.008	0.007			-0.027	-0.069			0.008	0.007
			(0.03)	(0.03)			(0.07)	(0.06)			(0.02)	(0.02)
MNE (NA)			0.052*	0.051*			0.095	0.079			0.052**	0.051**
			(0.03)	(0.03)			(0.07)	(0.07)			(0.02)	(0.02)
MNE (SA)			-0.033	-0.033			0.056	0.016			-0.033	-0.033
			(0.03)	(0.03)			(0.08)	(0.08)			(0.03)	(0.03)
MNE (AF)			0.031	0.03			0.031	0.090			0.031	0.030
			(0.04)	(0.04)			(0.11)	(0.10)			(0.03)	(0.03)
R^2	0.01	0.01	0.011	0.011	0.045	0.009	0.046	0.009	0.01	0.01	0.011	0.011
Observations:	5,847	5,847	5,847	5,847	5,847	5,847	5,847	5,847	5,847	5,847	5,847	5,847
Firm specifics:	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No

Note: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$ indicate the level of significance. Robust standard errors are used in the estimations. In the POLS column, the errors have been clustered by firm ID. The constant terms have been suppressed from the table output along with the year and industry dummy variables.

E Summary of the econometric methods for production functions estimation

Table 32 summarizes the introduced estimation approaches by several criteria: advantages, disadvantages, number of time periods required to perform the estimation and the possibility to account for exiters in panels. Such a representation is useful as it gives the reader a much clearer and structured overview of the trade-offs when different methods are selected. The Labor TFP is not included in the table as it does not involve any estimation or index construction.

Table 32: Comparison table of estimation methods

Productivity measure	OP_TFP	LP_TFP	I_TFP	OLS_TFP
Key advantages	Easy to incorporate exiters. Allows a nonlinear development of the error term	Relaxes some of the heavy OP requirements. Data on intermediate inputs is widely available and always positive	No econometric estimation and simple computation of the index. No requirements for the production function	Easy to compute in a statistical package. Takes account of all time-constant unobserved heterogeneity
Key disadvantages	Implausible assumptions on investment process. Identification issues in first stage estimation	Restrictions on the shape of the production function. Identification issues in first stage estimation	Assumptions for constant returns to scale and perfect competition are needed	Assumption that the productivity shocks are not correlated with the inputs
Min. t periods needed	2	2	1	2
Treatment of exiters	Possible	Possible	Possible	Possible

Note: OP is an abbreviation for the method proposed by Olley and Pakes (1996); LP for Levinson and Petrin (2003); I for the index methodology of Caves et. al (1982) and OLS for fixed effects transformation and estimation by OLS.