Foreign Investment and Domestic Productivity: Identifying Knowledge Spillovers and Competition Effects*

Christian Fons-Rosen, Şebnem Kalemli-Özcan, Bent E. Sørensen,
Carolina Villegas-Sanchez, and Vadym Volosovych

INCOMPLETE DRAFT: February 2017

Abstract

We identify knowledge spillovers from foreign investment to domestic firms using novel measures of "closeness" of foreign-owned and domestic and domestic firms in product space and in technology space. We rely on a new data set that spans six advanced countries connecting firms internationally in order to, a) separate competition effects on domestic firms from knowledge spillovers when domestic and foreign-owned firms are close in product space, and b) identify spillovers from foreign-owned firms that are close to domestic firms in technology space. We find strong negative competition effects on domestic firms that produce in the same four-digit sector as the foreign firms and positive knowledge spillovers to domestic firms operating in the same two-digit sector (but in *different* four-digit sectors). Using a measure of "technological closeness," building on the work of Bloom, Schankerman and van Reenen (2013), we find significant knowledge spillovers to firms which are close in technology space. On average, knowledge spillovers explain 60 percent of the total factor productivity improvement for domestic firms that are technologically close to foreign firms.

JEL: E32, F15, F36, O16.

Keywords: Multinationals, Competition, Technology, Selection, FDI, TFP.

*Affiliations: Universitat Pompeu Fabra and Barcelona Graduate School of Economics; University of Maryland, CEPR, and NBER; University of Houston and CEPR; ESADE - Universitat Ramon Llull; Erasmus University Rotterdam, Tinbergen Institute, and Erasmus Research Institute of Management. Carolina Villegas-Sanchez acknowledges financial support from Banco Sabadell.

1 Introduction

Many theoretical studies have argued that the presence of the multinational companies (MNCs) in an economy affects the productivity of purely domestic firms because of "spillovers." The empirical evidence is less clear. Several papers have found robust evidence that multinationals create positive "vertical spillovers" from customer-supplier relationships between firms in vertically linked sectors, that have input-output relationships. However, the evidence supporting positive "horizontal spillovers" that is, spillovers to domestic firms operating in the same sector as the MNC, is mixed.[1] The presence of the MNCs can lead to positive horizontal knowledge spillovers, through learning about technology, mimicking of production or sales practices, and/or mobile workers bringing skills acquired in foreign affiliates to domestic firms. It is hard to have direct measures of such activities, so researchers rely on reduced form regressions of the productivity of domestic firms on measures of the prevalence of the MNCs in the same sector of the domestic firm. Knowledge spillovers are likely to be contaminated by "competition effects" in such reduced form regressions, which may work through business stealing that will lead to lower sales or exit of domestic firms, or through a price effect, where domestic firms had to charge higher mark-ups to stay profitable or face higher input costs due to increased demand for inputs by the MNC.

We investigate the horizontal spillovers from foreign direct investment on domestic firms' total factor productivity (TFP) using novel measures of "closeness" between foreign-owned firms and domestic firms both in the product space and in the technology space. We perform reduced-form regressions of domestic firms' TFP on the MNC presence, but advance on previous work by defining carefully the domain where foreign and domestic firms interact and also by refining the measures of the MNC presence. We substantiate the TFP results by using data on patents to investigate whether domestic firms respond to the MNC presence by doing more R&D which will result in an increased number of patents.

It is standard to measure the MNC presence as the share of output produced by the MNC in a given sector. This measure is correlated with the dynamics of the sector growth. The MNCs often cluster in particular sectors as a result of, for example, endowments or sector-specific costs of transportation (medicine is usually not bulky and easily transported, while the opposite is true for building materials). If the transportation equipment sector in a country grows faster than the average sector, productivity may be high at the same time as the MNCs are attracted, or if the basic metal sector grows slower than the average sector productivity may be low at the same time as the MNCs stay away. Our dataset spans a number of countries and because our regressors are at the sector $\times$ country $\times$ year level, we control for sector $\times$ year fixed effects. Similarly, due to changes in economic policy or terms-of-trade, countries

[1] See Javorcik (2004) for evidence on vertical spillovers and see Keller (2004) for a summary of mixed results on horizontal spillovers and the literature in general. Specifically, Haskel, Pereira and Slaughter (2007) and Keller and Yeaple (2009) find positive horizontal spillovers in the UK and the United States, respectively; while Aitken and Harrison (1999) finds negative horizontal spillovers in Venezuela and Javorcik (2004) and Barrios, Gorg and Strobl (2011) find no horizontal spillovers in Lithuania and Ireland, respectively.

may have low productivity and be have few MNCs but, again, we control for such factors using country $\times$ year fixed effects. After including these fixed effects, no results are driven by different time-patterns in productivity between sectors or between countries.

It is obvious that competition from the MNCs will affect domestic firms producing "similar" goods. Our data and methodology will allow us to differentiate between firms producing similar goods in the same very narrow sector (four-digit) and firms producing similar goods in the same broader sector (two-digit). The production of similar, but not identical, products often relies on the same skills and, therefore, knowledge spillovers may be found on firms in the same broader sector as the MNC, when direct competition effects in the narrow sector are measured and taken out. In the literature, the typical "horizontal spillover regression" regresses the productivity of domestic firms on the presence of the MNCs in the same two-digit sector. As two-digit sectors are quite broad, this approach may not fully disentangle knowledge spillovers from competition effects. Using a new and a very large data sets that is not only representative in terms of registered domestic firms in each of the countries we use, but also represents international linkages for these firms in quite a compelling way, we can verify separately the effects of the MNC presence in the narrow four-digit sector and the broader two-digit sector.

Our approach allows us to contribute to the literature in several ways. First, we break down the traditional two-digit industry "horizontal spillover" into competition effects and "knowledge spillovers." We measure "knowledge spillovers" as being functions of the share of foreign firms' output operating within the same two-digit sector as the domestic firm but outside the same four-digit sector. If domestic and foreign firms operate in the same two-digit sector, but in different four-digit sectors, they are not direct product market competitors and the domestic firm can benefit from the technology advantage of the foreign firm. Second, building on Bloom, Schankerman and van Reenen (2013), who propose a method for estimating technology spillovers from research and development (R&D) activity. We examine if knowledge spillovers are particularly large when MNCs are present in sectors that are close in the "technology space"—measured from the similarity of patenting for pairs of industries.[2] Third, our multi-country, multi sector, firm level data set allows us to control for differential growth rates across sectors; i.e., sector-selection, with the use of two-digit sector-year fixed effects. Firm-level data are increasingly used in empirical research, but few sources exist for nationally representative data from several countries. And even fewer sources exist which connect firms internationally. We use an extensive large-scale dataset from Bureau van Dijk (BvD) on domestic firms and foreign firms, which is superior to alternative sources for our purpose.

We find significant intuitive results on TFP of domestic firms as a result of the MNC presence in the same two-digit broad sector of the domestic firm: competition spillovers from MNCs are negative while

[2]Our work is also related to work on competition and innovation and R&D; see for example Aghion et al. (2005), Bloom, Schankerman and van Reenen (2013), Jaffe (1988), Aghion and Howitt (1992).

knowledge spillovers from MNCs that are close in technology space are positive and highly significant. We show similar result showing patents instead of TFP. This is important since we use measured TFP, that is revenue TFP, where physical output and value of that output cannot be separated from each other. Therefore, it is hard to tell apart physical productivity improvements from price effects. This problem is shared by virtually the entire productivity literature and this issue is potentially important when estimating the effect of foreign direct investment on firm productivity, because markups may respond endogenously to competition effects, so that output prices charged by domestic firms rise, measured as a higher TFP and hence delivering a positive sign on the estimated spillover coefficient. We follow De Loecker and Warzynski (2012) and compute firm level markups. We show that 40 percent of the change in TFP induced by increases in foreign investment in technologically close sectors is driven by higher markups, with the remainder being changes in real productivity. We interpret these TFP improvements as resulting from explicit or implicit technology transfers.

We show that our results are quite distinct from vertical spillovers which are due to trade between firms which operate in different sectors. We calculate vertical spillovers measures using input-output matrices, replicate existing results in the literature, and demonstrate that horizontal spillovers are separate from vertical spillovers with both types will identified in our regressions. In terms of economic significance, horizontal spillovers are more important for TFP than vertical spillovers.

The rest of the paper is structured as follows. Section 2 lays out a conceptual framework. Section 3 presents our data. Section 4 discusses our measures and empirical framework. Section 5 presents benchmark results. Section **??** shows the robustness analysis. Finally, section 6 concludes.

2 Conceptual Framework

A large literature has modeled and documented that the most productive firms in a country are likely to become MNCs.[3] The presence of such firms in an economy is likely to affect domestic firms in various ways because the MNCs are likely to be more productive than the domestic firms in the host country. For example, (n.d.a) provide evidence on the productivity advantage of foreign firms and (n.d.b) document, using Spanish survey data, that formerly domestic firms indeed innovate, when becoming foreign owned.

We focus on the effect of the MNC entry on the productivity of domestic firms. First, the MNCs provide formidable competition for domestic firms. As the MNCs gain market share, domestic producers correspondingly lose market share and typically suffer a negative impact on productivity due to the smaller scale of production. However, some domestic firms will actively "fight back" and upgrade their technology in order to be more competitive. This induced effect is likely to be heterogenous across firms; for example, Aghion et al. (2005) explain how competition induces relatively more innovation in firms

[3]See Helpman (2006) for a survey.

that a priori are competitive while Bloom et al. (2016) document how increased competition from Chinese imports lead to innovations, as measured by patents. Second, the superior technology of the MNCs may directly impact the productivity of domestic firms through a number of more "passive" channels; for example, from copying management practices, hiring former employees of the MNCs, obtaining higher quality inputs, or being enticed to, or instructed in, providing better output for the MNC customers, or through various general equilibrium effects.

Reduced form estimates of productivity effects on domestic firms from the MNC entry is interpreted as learning effects. "Learning" typically is not directly observable, so this interpretation will be substantially supported if the entrance of the MNCs has disproportionate positive spillovers to firms that use similar technology. Bloom, Schankerman and van Reenen (2013) construct a measure of proximity for individual (large) firms using patenting data for 715 U.S. firms: firms that patent in similar categories are considered close in technology space. We build on this pathbreaking work and use their data to construct measures of similarity between four-digit sectors in Europe. Because these measures are based on technology, they are likely to be valid for advanced European economies, despite being constructed from U.S. data. In addition using U.S data is better for endogeneity concerns, where patenting in U.S. is exogenous to European firms.

Firms that produce similar products will be competitors, and while we do not observe the menu of goods produced by individual establishments, it is likely that firms belonging to the same four-digit sector will be competitors. Such firms may of course also use similar technologies creating spillovers for each other. We attempt to capture each of these mechanisms by constructing two measures of MNCs presence in the four-digit sector. One simply measures the presence of the MNCs in the four-digit sector and hence will help us to identify the competition effects. The other measures the presence of the MNCs in the four-digit sector that are also likely to use similar technology, allowing us to detect any knowledge spillovers even for firms who are close in product space operating in the same four-digit sector. Having such two measures, Ordinary Least Squares (OLS) estimates will capture the contribution of each which is orthogonal to the other, so the measure of the MNC presence is likely to capture pure competition effects while the measure of technologically close MNC presence is likely to capture the effect of knowledge spillovers.

What about firms which are in the same two-digit sector, but are not in the same four-digit sector? These firms are not likely to be directly competing for many goods; however, the presence of the MNCs in a given four-digit sector may still lead to knowledge spillovers to domestic firms that are in the other four-digit sectors different than MNCs'. This is because firms within a two-digit sector will produce similar, if not identical goods, given that they belong to different four-digit sectors. Because our measures of closeness in technology space measure how close the four-digit sectors are to each other within a given two-digit sector, we can examine if knowledge spillovers within two-digit sectors are particularly large if the MNCs enter four-digit sectors that are close to a given firm in the technology space, even MNCs are in different four-digit sectors than that of the domestic firms. As a result we will interpret this as prima facie evidence that knowledge spillovers involve actual technology transfer.

One caveat remains: firms within the same two-digit sector might have buyer-seller relationships through input-output relations between four-digit sectors and hence productivity effects from the MNC presence in the two-digit sector may capture vertical spillovers identified by Javorcik (2004) and others. Javorcik (2004) suggested that positive spillovers from FDI were more likely to be vertical than horizontal in nature, because multinationals have an incentive to prevent information leakage that would enhance the performance of their local competitors, but at the same time may benefit from transferring knowledge to their local suppliers. Because of the large size of our dataset, we can directly explore this issue using four-digit input-output tables to construct measures of the MNC presence in sectors which supply or demand goods from the sector of the domestic firm under analysis.

Direct evidence on innovation activity can be obtained by looking at patents and we use the number of patents granted to firms to explore the impact of foreign investment on the innovation activity of domestic firms. We find that domestic firms patent more when MNCs are close in technology space which provides important additional evidence on innovation in domestic firms in response to FDI in related technologically close sectors.

2.1 An Example

To fix ideas, we provide an example from our dataset. We choose Germany as the host country and analyze the domestic German company Tepper Aufzuege GMBH (www.tepper-aufzuege.de). Tepper manufactures elevators and it is classified in the four-digit subsector 2822 "Manufacture of lifting and handling equipment" that, in turn, belongs to the two-digit sector 28 "Manufacture of machinery and equipment, n.e.c." Also classified in subsector 2822, is the foreign company Loedige Aufzugstechnik GMBH (www.loedige-aufzuege.com) owned by KONE, the Finnish global leader in the elevator and escalator industry. Tepper and Loedige operate within the same four-digit industry and are potential direct competitors.

Operating in the same two-digit sector 28 as Tepper and Loedige is the Japanese subsidiary Kubota Baumaschinen GMBH (www.kubota-baumaschinen.de), which is classified in the different four-digit sector 2892 "Manufacture of machinery for mining, quarrying and construction." The traditional horizontal spillover measure used in the literature aggregates firms such as Kubota with competitors of the domestic company Tepper; however, these firms operate in different product markets. Kubota's sector 2892 supplies less than 1% of inputs used by Tepper's sector 2822 according to the four-digit Input-Output matrix. Conversely, sector 2822 supplies 4.5% of inputs used by sector 2892. Therefore, there is virtually no customer/supplier relationship between these two sectors. In addition, we searched for other evidence of the commercial relationship between Tepper and Kubota did not find any.

Nevertheless, Kubota produces excavators and Tepper may benefit from, for example, hydraulic technology advances achieved by Kubota, because the products of both companies (hydraulic excavators and hydraulic elevators) share a key component; namely, hydraulic cylinders and hydraulic motors.[4] Therefore, Tepper may benefits from knowledge spillover from Kubota. Empirically, our methods will capture this if Kubota contributes to our "knowledge spillover" variable for Tepper. In order to assess if sector 2822 is close to sector 2892 technology space, we use American firm-level patent data from the U.S. Patent and Trademark Office, provided by Bloom et al. (2013). 18% of patents registered by firms in NACE sector 2822, or 20% of the patent technology classes in Tepper's sector 2822, are similar to those of the patents of Kubota's sector 2892. Although these two sectors are not unusually close in technology space (the 18% number is around the 15th percentile in the distribution across all sector pairs), these sectors are much closer in technology space than in the product space. In our empirical analysis, we further follow Bloom et al. (2013) and use a measure adjusted for the firm R&D intensity, which captures that knowledge spillovers typically emanate from the technology leaders with extensive R&D operations.

The main take-away from this example is that horizontal spillover measures that are based on a sector classifications captures closeness in product product space but not closeness in technology space. Measures based on two-digit classifications, further confound the effect of direct product competitors (such as Loedige) and the effect of non-competitors, such as Kubota. By differentiating between competition and knowledge spillovers, between closeness in product and technology spaces, we obtain more precise estimates and separate our effects that may counteract each other.

3 Data

We use the ORBIS Global database from BvD supplemented with data from AMADEUS, also from BvD, which covers only Europe. The database covers more than 200 countries and over 10 million firms. The unit of observation is for most countries the firm, although in some countries data is reported at the establishment/plant level. Firms can be linked to their domestic and foreign parents through unique ID numbers.

BvD collects data from surveys and, in particular, national business and tax registries, and harmonizes them into an internationally comparable format. The ORBIS database provides a consistent time series representative sample for all countries under our analysis for both private and public firms. The extensive

[4] The movement of a hydraulic excavator is accomplished through the use of hydraulic fluid, with hydraulic cylinders and hydraulic motors. Hydraulic elevators, as opposed to geared and gearless traction elevators, use an underground hydraulic cylinder that pumps hydraulic fluid to raise a cylindrical piston and lift the load in low level buildings with two to five floors.

coverage of smaller, non-listed, firms is important when one is interested in aggregate economic effects because large firms (those with +250 employees) account for less than one percent of total manufacturing output or employment in the European countries under analysis. Other private vendors, besides BvD, compile such datasets using such national sources—prominent examples are Thompson&Reuters and Dun&Bradstreet. For our purpose, the ORBIS database is preferred because of its extensive coverage of domestic firms, which are the object of study here. In the ORBIS data set, we can cover up to 70 percent of the real economy in many countries—see Kalemli-Ozcan, Sorensen, Villegas-Sanchez, Volosovych and Yesiltas (2015) even if the coverage and representativeness across countries will depend on the rules and regulations of the particular country in question; in particular, many countries do not collect financial data from small private firms.

The ORBIS database has been previously used in the literature (see for example, Klapper, Laeven and Rajan (2006), Bloom, Draca and Van Reenen (2016), and Gopinath, Kalemli-Ozcan, Karabarbounis and Villegas-Sanchez (2015); however, this paper is the first to use foreign ownership information together with information on domestic firms.[5]

We focus on a sample of six advanced European countries (Belgium, Spain, Finland, France, Italy, and Norway) for the years 1999–2008. For this sample, we have nationally representative samples of firms with comprehensive information on the variables needed to compute total factor revenue productivity in the manufacturing sector.[6] We exclude micro enterprises (those with less than ten employees according to the European Commission definition). Very small firms are less likely to obtain technology spillovers from foreign investors and the ORBIS data are less representative for such small firms. Our final sample of firms accounts for more than ninety percent of manufacturing output in the countries under study and it avoids potential concerns about differences in the size distributions across countries. A further advantage of focusing on firms with more than ten employees is that we can directly compare our results to previous studies based on manufacturing surveys because previous work has been based on samples that include firms with a minimum threshold employment level (i.e., plus 10, 20, 50, or 100 employees depending on the study).

We next describe the main firm level variables used in the analysis: firm revenue productivity, foreign ownership, and number of patents. And we describe the complementary databases used to construct the spillover variables defined in Section 4.1. More details on the cleaning process and firm level statistics

[5]This likely reflect that the ORBIS ownership module is particularly burdensome to use for generating consistent time series. Access to ORBIS online or using the most current disk access will only provide ownership information as of the current date. Therefore, to have information that varies over time, and avoid survivorship bias, it is necessary to download the ownership structure of each firm from earlier disk vintages—see Kalemli-Ozcan, Sorensen, Villegas-Sanchez, Volosovych and Yesiltas (2015) for a detailed explanation on how to construct nationally representative firm level financial data and ownership structures from the BvD products.

[6]The data for Norway are for the period 2000-2008 because the available data for 1990ies are less nationally representative.

are provided in the Appendix.

3.1 Calculating Firm-Level Productivity

Our unit of analysis is "domestic firms," defined as firms that do not have foreign owners throughout the sample. We assume that firm i's output is given by a Cobb-Douglas production function:

$$Y_{it} \equiv \mathcal{F}(A_{it}; I_{i,t}^{\mathcal{J}}) = A_{it} L_{it}^{\beta_\ell} K_{it}^{\beta_k} , \tag{1}$$

where firm output is a function ($\mathcal{F}(\cdot)$) of physical productivity (A_{it}) and firm inputs ($I_{i,t}^{\mathcal{J}}$), where $\mathcal{J} = L, K$. L_{it} is the labor input, K_{it} is the capital input, β_k is the elasticity of output with respect to capital, and β_ℓ is the elasticity of output with respect to labor. We measure firm nominal value added, $P_{it}Y_{it}$, as the difference between gross output (operating revenue) and materials. We do not observe prices at the firm level, and we measure real output, Y_{it}, as nominal value added divided by Eurostat two-digit industry price deflators.[7] Labor input, L_{it}, is the firm's wage bill (deflated by the same two-digit industry price deflator).[8] Finally, we measure the capital stock, K_{it} as the book value of fixed assets and we deflate this value with the price of investment goods.[9]

To obtain firm level productivity estimates, we estimate the following log-value added production function:

$$y_{it} = \beta_0 + \beta_\ell \ell_{it} + \beta_k k_{it} + \omega_{it} + \epsilon_{it} , \tag{2}$$

where y_{it} is the log of real output, ℓ_{it} is the logarithm of labor input, k_{it} is the logarithm of capital input, ω_{it} is the logarithm of physical productivity and ϵ_{it} represents production shocks that are not observable by firms before making their input decisions at time t. The main concern, when estimating output elasticities with respect to the inputs in equation (2), is whether the firm observes its own productivity ω_{it} at the time of making input choices. This would render input quantities endogenous to productivity and OLS estimates of β_ℓ and β_k would be inconsistent; in particular, Ackerberg, Caves and Frazer (2015) argue that if the flexible labor input is chosen as a function of unobserved productivity, the coefficient on labor input is not identified in the approaches proposed by OP and LP. We follow the approach suggested in Wooldridge (2009), which builds on previous work by Olley and Pakes (1996) (OP) and Levinsohn and

[7] Norway and France do not have good coverage of industry price deflators at the two-digit level and we use the total manufacturing industry price deflator for these two countries.

[8] Using the wage bill, rather than the head count, helps adjust for differences in the quality of the workforce across firms because more skilled workers normally are paid more.

[9] We do not have industry-specific investment deflators and, therefore, we use country-specific prices of investment from the World Development Indicators to deflate the book value of fixed assets. The capital stock includes both tangible and intangible assets because in 2007 there was a change in the accounting system in Spain and leasing items that until 2007 had been part of intangible fixed assets were from 2008 included under tangible fixed assets. To avoid breaks in the capital stock time series we opt to use the sum of tangible and intangible fixed assets as our measure of capital stock. Results are robust to estimating TFPR using tangible fixed assets as a measure of capital in the other countries of the sample.

Petrin (2003) (LP), that addresses the concerns raised by Ackerberg, Caves and Frazer (2015).[10]

We estimate the production function by country and two-digit sector (Table B.1 in Appendix B shows the estimated elasticities) and winsorize the resulting distribution at the 1 and 99 percentiles by country. Our main results are qualitatively robust to estimating productivity from a translog production function following the methodology proposed in De Loecker and Warzynski (2012).

3.2 Firm-Level Foreign Ownership

The ownership section of ORBIS contains detailed information on owners of both listed and private firms, including name, country of residence, and type (e.g., bank, industrial company, private equity, individual). The database refers to each record of ownership as an "ownership link." An ownership link indicating that an entity A owns a certain percentage of firm B is referred to as a "direct" ownership link. BvD traces a direct link between two entities even when the ownership percentage is very small (sometimes less than one percent). For listed companies, very small stock holders are typically unknown.[11] We compute *Foreign Ownership* as the sum of all percentages of *direct* ownership by foreigners.[12] We define a firm to be "domestic" only if it never had *any* type of foreign owner during the sample period. Figure 1 shows how foreign-owned firms are over-represented in high-technology sectors and Figure 2 shows how foreign-owned firms prefer controlling stakes in high-technology sectors.

3.3 Firm-Level Patents

We match the patent data provided by Bloom, Draca and Van Reenen (2016) to our sample of firms. Bloom, Draca and Van Reenen (2016) identify the firms in AMADEUS that filed at least one European Patent Office (EPO) patent since 1978 and provide the number of granted patents by firm. Patents are dated by application year in order to measure the formal invention year of the patent (see Bloom, Draca and Van Reenen (2016) for more details on the matching procedure between AMADEUS and the EPO dataset). Using the unique BvD firm identifier, we match our sample of firms to that of Bloom, Draca and Van Reenen (2016). We assume that unmatched firms were not granted any patents. We observe the number of granted patents, but we do not have information on the characteristics of each individual

[10] See Appendix B for a detailed description of the estimation procedure.

[11] Countries have different rules for when the identity of a minority owner needs to be disclosed. France, Germany, the Netherlands, and Sweden demand that listed firms disclose all owners with more than a five percent stake, while disclosure is required at three percent in the UK, and at two percent in Italy. BvD collects ownership data from official registers (including SEC filings and stock exchanges), annual reports, private correspondence, telephone research, company websites, and news wires.

[12] For example, if a company has three foreign owners with stakes of 10, 15, and 35 percent, the foreign ownership fraction for this company is 60 percent.

patent, such as the patent activity cluster, and we do not have information about R&D expenditures at the firm level.

3.4 Sector Technology Closeness

In order to construct measure of closeness between sectors in technology space, we utilize the technology closeness measures provided by Bloom, Schankerman and van Reenen (2013) for pairs of firms. They use U.S. firm level accounting data (sales, employment, capital, R&D expenditure etc.) and market value data from Compustat over the period 1980–2001 in order to match firms to U.S. Patent and Trademark Office data from the NBER data archive. The matching results in an unbalanced panel of 715 listed firms with at least four observations between 1980 and 2001. They assign patents to different technology clusters and compute a technology closeness measure, which is calculated as the uncentered correlation between all firm i, j pairings. We use information on technology closeness between firm pairs and information on firms' R&D expenditures to compute bilateral technology closeness between four-digit sectors (see the next section for a detailed description on how these measures are constructed). The number of four-digit industries covered in Bloom, Schankerman and van Reenen (2013) is lower than the number of four-digit industries available in our dataset and, after we drop sectors with no counterpart in Bloom, Schankerman and van Reenen (2013), we end up with a final sample of 570,000+ firm-year observations.[13] Table C.1 in Appendix XXX shows that there are no major differences in company size across the original sample and the merged sample. Because the sectors available in Bloom, Schankerman and van Reenen (2013) is subset of ours, the number of sectors available is dictated by the number of sectors available that. However, this does not imply that we select only sectors with positive technology links. Once we aggregate the dataset in Bloom, Schankerman and van Reenen (2013) to the sector level, there can be sectors with no technological links. Table C.2 in Appendix XXX reports summary statistics for the main variables and the final sample of firms used in the analysis. Table C.5 shows the correlation XXX what correlation XXX.

3.5 Sector Vertical Linkages

The vertical spillovers literature is based on the idea that the productivity of domestic firms supplying intermediate inputs to foreign owned firms can be fostered if the MNC transfers soft and/or hard technology to their suppliers. Ideally, we would like to explore this channel using firm level information on the productivity of domestic firms that start supplying foreign owned firms. However, firm-to-firm information on sales and purchases of inputs is not available in most datasets and instead, sector input-

[13]The original dataset covers 299 four-digit industries while the merged dataset covers 135 four-digit industries.

output tables are used to proxy input flows across industries. In our study, it is crucial to identify vertical linkages at the four-digit industry level and unfortunately, there is no such information for the wide set of countries in this analysis.[14] We therefore use the U.S. input-output table from the Bureau of Economic Analysis (BEA).

The use of the United States as a benchmark has a long tradition in empirical economics, see for example Rajan and Zingales (1998) and Acemoglu, Johnson and Mitton (2009) for prominent contributions, and we follow that approach for constructing sectoral measures of technology closeness and vertical integration. This implicitly assumes that the six advanced European countries of our sample are close to the United States in terms of technology, but if the U.S. Input-Output structure and U.S. sector technology closeness are imperfect for European countries, we are introducing random error in the measurement of our regressors and, therefore, reducing the probability of finding statistically significant results; in other words, it is highly unlikely that this approach results in spuriously significant results. An advantage of using U.S. data, is that we avoid possible endogeneity of the constructed regressors. Further, we use time-invariant measures of the technology closeness that pre-date the sample of analysis in order to avoid potential endogeneity of regressors. And, finally, identification does not require that the technology closeness measures of European sectors are exactly the same as those of the United States, but rather that the ranking of sectors remains stable across countries, which seems plausible.

4 Empirical Methodology

4.1 Calculating FDI Spillovers

Horizontal Spillovers. The typical horizontal spillover relation estimated in literature, see for instance Javorcik (2004), computes exposure to foreign owned firms as the share of foreign output in total two-digit industry output:

$$\textsc{horizontal}_{s2,t} = \frac{\sum_{i\in s2} \mathrm{fo}_{i,t} \times \mathrm{go}_{i,t}}{\sum_{i\in s2}, \mathrm{go}_{i,t}}, \tag{3}$$

where $s2$ refers to the two-digit sector classification, $\mathrm{go}_{i,t}$ refers to gross output (operating revenue) of firm i at time t and fo is the percentage of foreign-owned capital of firm i.

Non-Technology Spillover Measures. The industry classification used in equation (3) replicates that of many empirical papers. We propose two-digit HORIZONTAL measures based on four-digit sector infor-

[14] The World Input-Output Database (WIOD) provides time-series of world input-output tables for forty countries worldwide however, only at the two-digit industry level.

mation:

$$\textsc{comp}_{s4,t} = \frac{\sum_{i \in s4} \text{fo}_{i,t} \times \text{go}_{i,t}}{\sum_{i \in s4} \text{go}_{i,t}} \quad \text{and} \quad \textsc{know}_{s4,t} = \textsc{horizontal}_{s2,t} - \frac{\sum_{i \in s4} \text{fo}_{i,t} \times \text{go}_{i,t}}{\sum_{i \in s2} \text{go}_{i,t}}\,, \tag{4}$$

where $s4$ indicates a four-digit sector classification and the rest of the notation is identical to that in equation (3). Exploiting information on the four-digit industry classification of firms, these two equations break the aggregate $\textsc{horizontal}_{s2,t}$ variable into two components: $\textsc{comp}_{s4,t}$ which captures the "competition" effect (i.e., the effect of foreign-owned firms producing in the same four-digit industry as the domestic firm) and the $\textsc{know}_{s4,t}$ or the knowledge spillover that measures foreign presence in the same two-digit sector, *excluding* output produced by foreign-owned companies in the same four-digit sector. This variable is intended to capture the relationship between foreign-owned and domestic firms producing within the same two-digit sector but excluding the direct competitors. For example, in sector 29 "Manufacture of motor vehicles, trailers and semi-trailers," we would expect positive interaction effects between a foreign-owned car manufacturer (four-digit sector 29.10 "Manufacture of motor vehicles") and domestic owned manufacturers of electrical and electronic equipment for motor vehicles (sector 29.31).

Technology Spillover Measures.

We construct measures of MNC presence by technological proximity. To obtain proxies of sectors' technology closeness, we rely on the work by Bloom, Schankerman and van Reenen (2013). They construct firm-level measures of technological closeness by measuring the similarity of patenting by using data from U.S. Patent and Trademark Office. The patent data are categorized by 426 different technology classes. Bloom et al. (2013) calculate the average share of patents each firm holds in each such technology class over the period 1970–1999. They then define for each firm i the vector of i's technological activity $\text{t}_i = (\text{t}_{i1}, \text{t}_{i2}, ..., \text{t}_{i426})$, where t_{ix} is the share of patents of firm i in technology class x. Then, for each firm pair i, j in the sample, Bloom et al. (2013) construct measures of technology closeness denoted tech_{ij}. Following Jaffe (1986), the measure of firm technology closeness is the uncentered correlation of patent share vectors t_i and t_j:

$$\text{tech}_{ij} = \frac{(\text{t}_i \text{t}_j')}{(\text{t}_i \text{t}_i')^{1/2} (\text{t}_j \text{t}_j')^{1/2}}\,. \tag{5}$$

Bloom et al. (2013) collect the firm-level R&D expenditures, labelled r_{jt} for firm j at time t, and test if a given firm i is affected by the R&D of other firms $j \neq i$ that are technologically close, as measured by tech_{ij}. Specifically, they construct a measure $\text{spilltech}_i = \Sigma_j \text{tech}_{ij} \text{r}_j$ and test if this empirically predicts the R&D expenditures and other outcomes of firm i. Bloom et al. (2013) uncover large effects of R&D on technologically close firms.

We use Bloom et al. (2013)'s firm-level measures of technology closeness but aggregate them to the sector level, matching to our NACE industrial classification. The following steps describe our procedure where we use their notation, except that we use a small-cap font for firm-level variables to contrast with the country-sector variables in capitals. First, for each four-digit sector pair, we compute the *sectoral* technological closeness as the R&D-weighted sum of the technology closeness of firms operating in sector pairs $s4$ and $\tilde{s}4$ as:

$$\text{SPILL\_RD}_{s4,\tilde{s}4} = \sum_{i \in s4} \sum_{j \in \tilde{s}4} \text{tech}_{ij} \times \left(\frac{\mathrm{r}_i}{\mathrm{r}_{s4} + \mathrm{r}_{\tilde{s}4}} + \frac{\mathrm{r}_j}{\mathrm{r}_{s4} + \mathrm{r}_{\tilde{s}4}} \right), \tag{6}$$

where $s4$ denotes a particular four-digit industry and $\tilde{s}4$ refers to four-digit industry sectors in the economy other than $s4$. The technology closeness measure and the firm level R&D expenditure are time invariant because we use the average R&D expenditure for each firm over time. Figure 3 shows the distribution of the sectoral technological closeness measure (6) split by the alternative broad OECD classification of the technology intensity of the sector. Sectors that are classified by the OECD as technologically intensive (represented by the solid green line) feature larger values of our technological closeness measure (6).

When we take into account both commonality of patents and sector sizes in terms of R&D spending the technological proximity of the two sectors featured in the example in Section 2 becomes even more prominent. Based on the measure in equation (6) the sector 2822's closeness to technological advances in sector 2892 falls into 95th percentile in the in-sample empirical distribution of these measures, even above the degree of technological closeness of sector 2822 to itself (that falls into approximately 92nd percentile).[15]

The effect of technology transfers from multinationals in a particular sector, $\tilde{s}4$, which is technologically linked to domestic firms in sector $s4$, may be small if sector $\tilde{s}4$ has low output. Therefore, we refine the knowledge spillover measure $\text{KNOW}_{s4,t}$ by introducing weights that reflect the economic importance (in terms of sectoral gross output, GO) of the four-digit sectors $\tilde{s}4$ that are technologically linked to a given four-digit sector $s4$, as:

$$\text{WTECH}_{s4,\tilde{s4},t} = \frac{\text{SPILL\_RD}_{s4,\tilde{s}4} \times \text{GO}_{\tilde{s}4,t}}{\sum\limits_{\substack{\hat{s}4 \in s2(s4) \\ \hat{s}4 \neq s4}} \text{SPILL\_RD}_{s4,\hat{s}4} \times \text{GO}_{\hat{s}4,t}}, \tag{7}$$

where $s2(s4)$ denote the 2-digit sector that includes the 4-digit sector $s4$. The denominator in equation (7) is chosen such the weights sum to unity in equation (8). The numerator is the technological similarity between sectors $s4$ and $\tilde{s4}$ weighted by relative research intensity and output. Our empirical analysis controls for factors affecting 2-digit sectors and the weights are therefore designed to reflect56x the relative technological closeness within 2-digit sectors. We define $\text{WTECH}_{s4,\tilde{s4},t}$ for $\tilde{s4}$ sectors in $s2(s4)$. and then

[15]The 90(95)th percentile of this measure in our sample is .087(.118) while the closeness of the receiver sector 2822 to supplier sector 2892 is .112 and to own supplier sector 2822 is .092.

use these weights to compute:

$$\text{KNOWTECH}_{s4,t} = \sum_{\substack{\tilde{s}4 \neq s4 \\ \tilde{s}4, s4 \in s2}} \text{WTECH}_{s4,\tilde{s4},t} \times \text{COMP}_{\tilde{s}4,t} \,. \tag{8}$$

The summation is done over four-digit industries $\tilde{s}4$, within the two-digit sector of $s4$, other than the given four-digit sector $s4$. The WTECH weights are normalized so that the values sum to unity for each sector $s4$ implying that an across-the-board increase in foreign ownership lead to an across-the-board increase in $\text{KNOWTECH}_{s4,t}$ of the same magnitude.

Similarly, we define the following weights for the four-digit sector $s4$:

$$\text{WTECH}_{s4,s4,t} = \frac{\text{SPILL\_RD}_{s4,s4} \times \text{GO}_{s4,t}}{\sum_{\hat{s}4 \in s2(s4)} \text{SPILL\_RD}_{s4,\hat{s}4} \times \text{GO}_{\hat{s}4,t}} \,, \tag{9}$$

and compute the following technology weighted measure for the within four-digit-sector spillover:

$$\text{COMPTECH}_{s4,t} = \text{WTECH}_{s4,s4,t} \times \text{COMP}_{s4,t} \,. \tag{10}$$

The COMPTECH measure interacts the presence of foreign firms in the four-digit sector $\text{COMP}_{s4,t}$ with a measure of the importance of technological linkages between firms within the sector. COMPTECH is obviously correlated with the COMP measure, but when we include both in a regression, COMPTECH will mainly capture the importance of closeness in the technology space while COMP will capture the competition effect from having similar principal output.

Figure 4 shows the relative importance of each type of spillovers according to the OECD classification of manufacturing industries into technological categories.[16] It appears that the importance of KNOWTECH (the blue bars) is higher in sectors classified as high-technology.

One might worry whether the variable KNOWTECH contains information much different from the variable KNOW (the red bars). Figure 5 in panel (a) shows that, while the correlation among the two variables is high, there is substantial variation in the variable KNOWTECH not captured by KNOW and vice versa.

Table C.2 in the appendix provides summary statistics for all variables used in the analysis.

[16]The OECD classification is based both on direct R&D intensity and R&D embodied in intermediate and investment goods. Four categories were introduced: high-, medium-high, medium-low and low technology. Figure C.1 in the appendix shows a more detailed breakdown by the two-digit industry.

4.2 Mapping to Regressions

Most existing studies of FDI spillovers has been limited to single countries and have typically estimated an equation of the following type for the sample of domestic firms:

$$\log\left(\mathrm{TFPR_{i,s2,t}}\right) = \beta\textsc{horizontal}_{s2,t-1} + \alpha_i + \delta_t + \epsilon_{i,s2,t}, \tag{11}$$

where $\mathrm{TFPR_{i,s2,t}}$ refers to revenue total factor productivity of firm i, in sector $s2$, at time t, where $\textsc{horizontal}_{s2,t}$ is a regressor which captures the presence of foreign ownership in (two-digit) sector $s2$. α_i represents firm-specific dummies and δ_t represents year dummies. The parameter of interest is β and a positive coefficient indicates positive productivity spillovers from foreign-owned companies to domestic firms. With firm-fixed effects included, β captures the correlation between the changes in the HORIZONTAL variable and changes in firm TFPR at the sector level over the year of the sample.

There are two concerns with this specification. The first concern is whether we are estimating a causal relation; i.e., where we can interpret β as the change in firm level TFPR derived from exogenous changes in foreign activity in a particular sector. However, FDI will be endogenous if the choice of foreign investment is attracted to sectors where productivity growth relatively fast or sectors where slow growth leads to low valuations, which then attracts investors.[17] An important advantage of a multi-country setting is that we estimate relations which control for patterns in sector-level productivity by the inclusion of sector-year fixed effects and we, therefore, estimate the following equation:

$$\log\left(\mathrm{TFPR_{i,s2,c,t}}\right) = \beta\textsc{horizontal}_{s2,c,t-1} + \alpha_i + \phi_{s2,t} + \delta_{c,t} + \epsilon_{i,s2,c,t}\,, \tag{12}$$

where $\mathrm{TFPR_{i,s2,c,t}}$ refers to total factor productivity of firm i, in sector $s2$, country c, at time t and the terms $\delta_{c,t}$ and $\phi_{s2,t}$ represent country-year and sectoral-year fixed effects, respectively.

The second concern related to specification (11) is, as we have emphasized, the coarse level of aggregation of the spillover variable based on the two-digit industry classification. To address this concern, we break down the HORIZONTAL spillover variable into COMP and KNOW and estimate the following equation:

$$\log\left(\mathrm{TFPR_{i,s4,c,t}}\right) = \beta_1\textsc{comp}_{s4,c,t-1} + \beta_2\textsc{know}_{s4,c,t-1} + \alpha_i + \phi_{s2,t} + \delta_{c,t} + \epsilon_{i,s4,c,t}\,, \tag{13}$$

where $\mathrm{TFPR_{i,s4,c,t}}$ refers to total factor productivity of firm i, in sector $s4$, country c, at time t and the terms $\delta_{c,t}$ and $\phi_{s2,t}$ represent country-year and sector-two-digit-year fixed effects, respectively. We expect β_2 to be positive, capturing the possible productivity improvement effects derived from the presence

[17] The direction of bias due to the endogeneity of FDI is not obvious. On the one hand, certain sectors may have high productivity growth and therefore attract foreign investors. On the other hand, foreigners may buy lemons because they have less information than domestic investors.

of foreign-owned firms which are not competitor but produce in closely related sectors. The dependent variable is revenue productivity and, therefore, a positive coefficient does not speak to whether the increase in firm revenue productivity is driven by technology transfer or demand-price effects. We abstract from this discussion for now and return to it in section 5.3.

Finally, we augment equation (13) with the technology-weighted knowledge spillover variable. The main specification estimated is the following:

$$
\begin{aligned}
\log\left(\mathrm{TFPR}_{\mathrm{i,s4,c,t}}\right) \quad = \quad & \beta_1 \mathrm{COMP}_{s4,c,t-1} + \beta_2 \mathrm{KNOW}_{s4,c,t-1} + \\
& \beta_3 \mathrm{COMPTECH}_{s4,c,t-1} + \beta_4 \mathrm{KNOWTECH}_{s4,c,t-1} + \\
& \alpha_i + \phi_{s2,t} + \delta_{c,t} + \epsilon_{i,s4,c,t}\,,
\end{aligned} \tag{14}
$$

where $\mathrm{TFPR}_{\mathrm{i,s4,c,t}}$ refers to total factor revenue productivity of firm i, in sector $s4$, country c, at time t and the terms $\delta_{c,t}$ and $\phi_{s2,t}$ represent country-year and sector-two-digit-year fixed effects, respectively. We expect β_1 to be negative once the potential positive effect from knowledge transfers by direct competitors is accounted for β_3, while β_4 will be positive if domestic firms benefit from the presence of foreign investors in closely related technological sectors.

5 Results

In order to relate to previous work, we start with the basic specification previously used in the literature which has focused on closeness in product space. In Table 1, we report the results from specification (11) using the spillover measure HORIZONTAL, which is defined as foreign ownership in the same two-digit sector as has been done in a number of works. Firms are heterogenous and while most of the existing literature estimates equations similar to equation (11) by Ordinary Least Squares (OLS), heteroskedasticity is of particular concern in the current study that pools firms across different countries with potentially different firm-level volatility. Therefore, all our results are estimated by two-step feasible Generalized Least Squares (GLS).[18]

Table 1 about here

[18]The first step estimates the equation by OLS, and for each firm the standard error of the innovation term is estimated as the square root of the mean (over the observations for the firm) squared residuals. In the second step, the least squares regression is repeated, weighting each firm by the inverse of its estimated residual standard error. GLS, although less so than OLS, can be sensitive to the effects of outliers and because the innovation standard errors are estimated from a low number of observations, they are somewhat imprecise. We, therefore, winsorize the lower tail of the weights distribution at 5 percent in order to avoid giving any individual firm excessively high weight. Graphical inspection of a partial correlation plot of the regression reveals that there are no obvious outliers in our second step regression. Similar results were found if the weights were obtained with a parametric model of the error variance (i.e., estimating standard errors as a function of firm characteristics).

We include country-year fixed effects (corresponding to time fixed effects in a single-country study) and firm fixed effects. Because our regressors do not have magnitudes which are immediately intuitive, we report standardized coefficients, which are the coefficients obtained when each regressor is normalized by its standard deviation (those standard deviations are reported at the bottom of the respective table). The benefit of reporting normalized residuals is that the magnitude of the reported parameter reflects the importance of the corresponding regressor in explaining productivity; i.e., the relative size of the parameter of a regressor reflects the relative importance of the regressor in explaining the dependent variable.

In the specification reported in column (1), spillovers are not significant, consistent with many previous paper. However, in column (2), where we control for factors that affect entire sectors in each given year, using fixed effects, we find that foreign presence in the two-digit sector increases the productivity of domestic firms. The estimated coefficients are semi-elasticities with the left-hand side scaled by 100; i.e., such that the estimated coefficients are interpreted as the predicted percentage change in the dependent variable following a unit change in the regressor. Therefore, the coefficient of 0.515 implies that an increase in foreign presence in the two-digit industry of one standard deviation increases domestic firms' productivity by about 0.5 percent. (The reader who wants to consider the effect of a given change in the regressor can use the standard deviations of the regressors reported at the bottom of the table, but we will postpone the discussion of economic magnitudes till we discuss Table 8.) In column (3), we use our finer sector-classification, estimating specification (13), which allows for a potential competition effect within the four-digit sector, captured by the variable COMP and learning from foreign presence in sectors with similar but not very similar output; i.e., presence in the two-digit sector but outside the four-digit sector, captured by the variable KNOW. We find a highly significant competition effect within the four-digit sector with a decline in productivity of 0.115 percent following a one standard-deviation change in the COMP-measure. The effect of changes in the KNOW-measure is positive as expected, with a large coefficient of 0.155, but it is less precisely estimated and it is significant only at the 1-percent level.

Next, we include measures of closeness in technology space. In Table 2, the first column is repeated from Table 1 for convenience, while the other columns explore whether, in specification (14), closeness in technology space is more important than closeness in product space for competition and knowledge spillovers. In column (2), we include the KNOWTECH measure and this measure is highly significant and renders the basic non-weighted measure insignificant. This is intuitive: firms learn from foreign firms with similar technology. The result may not be surprising, given that Bloom, Schankerman and van Reenen (2013) show that firms learn from the technological innovation of firms that are close in technology space, but the relevance for FDI spillovers seem not to have been realized before. The estimated coefficient to the standardized technology-weighted measure is highly significant with a value of 0.298, implying that a one standard deviation increase in the presence of foreign firms close in technology space (albeit outside the four-digit sector) increases productivity by about a third of a percent. An important, hitherto unappreciated, implication of this result is that the amount of knowledge spillovers that country enjoys from FDI depends not just on *how much* foreign investment there is, but also critically on *in which firms* foreign investment takes place. In particular, as we quantify below, the impact will be larger when FDI is concentrated in sectors where technological spillovers to other sectors are high, and lower if FDI is concentrated in "technologically isolated" sectors. This may well explain why the estimated productivity spillovers from FDI differ across countries. In other words, the productivity enhancing effects of FDI for local firms will depend on the interaction between high foreign inflows into high technological sectors that are close in technology space.

The COMP measure, in column (2), is unchanged from column (1), which indicates that it unsurprisingly is uncorrelated with the KNOWTECH measure. In column (3), we include the COMPTECH measure of technology-weighted foreign presence in the same four-digit sector—this measure is a priori hard to sign: firms in the same four-digit sector are likely to be competitors but learning may also be intense from competitors which are close in technology space. We have measures of four-digit closeness in both product- and technology space and intuitively the unweighted COMP measure will mainly capture competition effects while the COMPTECH measure is likely to capture knowledge spillovers after controlling for COMP. And, indeed, the latter measure takes a positive highly significant coefficient of 0.644, while the coefficient to COMP takes a highly significant value of –0.508, consistent with increased competition hurting domestic firms as previously found in the previous table, but the smaller coefficient found there reflects that foreign firms producing in the same four-digit sector delivers both negative competition effects and positive knowledge effects and only after constructing our new measures can we teach out those opposing effects. The KNOW regressor remains insignificant, indicating that closeness in product space in the two-digit sector is not important when closeness in technology space in the two-digit sector has been controlled for (and within four-digit sector effects are controlled for). Column (4) simply verifies that dropping the variable KNOW has no substantial effect on the other estimated coefficients.

Table 2 about here

Table 2 explains the conflicting results found so far in the horizontal spillover literature, showing evidence of positive spillovers at the two-digit industry level in developed countries, negative effects in emerging markets or no effects in both developed and emerging markets.

5.1 Predicted Effects of Increased Foreign Ownership and Aggregate Implications

The aim of this section is to provide evidence on the goodness of fit of the model and *quantify* the average TFPR effects implied by changes in foreign ownership. First, Figure 6 shows the correlation between changes in actual log TFPR and changes in the predicted values of log TFPR. To compute the actual change in log TFPR between year 2000 and 2008, we first calculate the average TFPR level in each country, four-digit sector, year triplet based on the firm level information. We then compute the growth rate for each country-four-digit sector as the log difference between year t and year $t-1$. Finally, the overall TFPR growth rate is the weighted average of the country-four-digit sector growth rates. We follow the same steps to compute the average predicted TFPR growth rate during the period 2000-2008 (see the detailed steps in the next subsection). The figure shows that the model can explain 43 percent of the variation in sectoral TFPR.

The second aim of this section is to quantify the average effects implied by changes in foreign ownership according to the effects estimated in column (4) of Table 2. We consider the effect of an across-the-board change in fo_i, of magnitude Δfo, which should be considered as widely distributed across firms. That is, for all domestic firms, we use the estimated equation (14) to predict a change of Δfo for all firms.[19]

The predicted change in firm-level TFPR is:[20]

$$\begin{aligned} \Delta \log\left(\text{TFPR}_{\text{i,s4,c}}\right) &= \hat{\beta}_1 \Delta\text{COMP}_{s4} + \hat{\beta}_3 \Delta\text{COMPTECH}_{s4} \\ &\quad + \hat{\beta}_4 \Delta\text{KNOWTECH}_{s4} \end{aligned} \tag{15}$$

All right-hand-side variables in equation (15) are sector level variables, which implies that we obtain the predicted average sectoral TFPR increase. In what follows, we derive analytical expressions for the change in the right hand side variables associated with an across-the-board increase in foreign ownership (Δfo). The change in the COMP variable is:

$$\Delta\text{COMP}_{s4} = \frac{\sum_{i \in s4} \Delta\text{fo} \times \text{go}_i}{\sum_{i \in s4} \text{go}_i} = \Delta\text{fo}. \tag{16}$$

[19] Literally, this means that some firms will have more than 100 percent foreign ownership, adjusting for this in simple add hoc ways would not change the predicted effects.

[20] Notice the effect of β_2 is suppressed from this specification because the results in Table 2 show that the variables KNOW and KNOWTECH are both proxies for technology transfer within the same two-digit industry (excluding the effect of foreign-owned firms within the same four-digit sector) and the variable KNOW is no longer statistically significant once the effect of KNOWTECH is accounted for.

the change in COMPTECH will be:

$$\Delta\text{COMPTECH}_{s4} = \text{WTECH}_{s4,s4} \times \Delta\text{fo}. \tag{17}$$

and finally, the change in KNOWTECH will be:

$$\Delta\text{KNOWTECH}_{s4} = \sum_{\substack{\tilde{s4} \neq s4 \\ \tilde{s4}, s4 \in s2}} \text{WTECH}_{s4,\tilde{s4}} \times \Delta\text{COMP}_{\tilde{s4},t} = \sum_{\substack{\tilde{s4} \neq s4 \\ \tilde{s4}, s4 \in s2}} \text{WTECH}_{s4,\tilde{s4}} \times \Delta\text{fo} = \Delta\text{fo}\,. \tag{18}$$

Recall that the weights in equation (18) are constructed such that $\sum_{\substack{\tilde{s4} \neq s4 \\ \tilde{s4}, s4 \in s2}} \text{WTECH}_{s4,\tilde{s4}} = 1$ and therefore, the last equality in equation (18) follows.

Collecting terms, the total effect of an across-the-board change ($\Delta\text{fo}_i = \Delta\text{fo}$) is

$$\Delta \log\left(\text{TFPR}_{\text{s4,c}}\right) = \left(\hat{\beta}_1 + \hat{\beta}_3\,\text{WTECH}_{s4,s4} + \hat{\beta}_4\right)\Delta\text{fo}\,. \tag{19}$$

We follow equation (19) to obtain the average change in log TFPR at the four-digit sector and finally, compute the total change in aggregate country level TFPR as the weighted average of the four-digit sectoral predicted changes.[21]

FDI may not take place across the board. As is clear from equations ((18)) and ((17)), the impact of FDI is higher on sectors where the WTECH is higher. For each sector $s4$, we can calculate the "average connectedness" as

$$\sum_{\substack{\tilde{s4} \neq s4 \\ \tilde{s4}, s4 \in s2}} \text{WTECH}_{s4,\tilde{s4},t} \tag{20}$$

In Table 9, we tabulate predicted effects. Our variables, except the technology weighted competition measure, COMPTECH, are scaled to be proportional to the change in foreign ownership, so increasing FDI by a similar amount for all firms, will have an effect that roughly can be predicted from the sum of the coefficients. More precisely, an increase in foreign ownership of 10 percentage points will lead to predicted increase in productivity of 0.42 percent. This effect is the sum of three partial effects, a 0.20 percent increase due to spillovers to firms that are close in technology space and not competing because they are outside the four-digit sector. A 0.54 percent increase from knowledge spillovers from firms in the same four-digit sector which is off-set by competition effect of –0.32 percent.

A very important implication of our results is that the spillover effects depend on the location of FDI. TO COME.

[21] Weights are computed as the share of four-digit sectoral output in total manufacturing output in each country and year. To avoid results being driven by changes in output over time, we normalize the weights to add up to one by computing the average four-digit weight across time and re-scaling by the sum across all four-digit time-invariant weights in the economy.

5.2 Direct Evidence on Technology Spillovers: FDI and Domestic Firm Patenting

Our spillover regressions provide clear evidence that an increasing presence of foreign firms impacts significantly on the productivity of domestic firms. Such regressions do not inform about whether these effects reflect scale economies, passive learning, or active innovation by the domestic firms exposed to foreign presence. However, the firm-level patent data for Europe allow us to dig somewhat deeper by examining if patenting activity changes with with foreign exposure. We regress the logarithm of one plus the number of patents on our spillover measures.[22]

Table 3 about here

Table 3 presents the results for patenting.[23] The results are easily summarized: foreign presence in the same four-digit sector has no significant effect, although the point estimate for the non-technology weighted measure is negative and it is possible that there are heterogenous effects within four-digit sectors. The impact of technologically close foreign firms in the four digit sector is highly significant with a one-standard error increase in our measure leading to roughly a 0.15 percent increase in patenting.[24] In column (1), foreign presence in the product space outside of direct competitors implies higher patenting, but when we include the technology weighted measure in column (2), this technology weighted measure is highly significant and the non-technology weighted measure becomes insignificant. Further including the technology weighted measure of foreign presence in the four-digit industry delivers an insignificant coefficient and has no impact on the estimated coefficients of the other measures.

Our interpretation of these results is that the presence of foreign firms using similar technology to a given domestic firm creates opportunities the domestic firm to actively innovate, as witnessed by the higher patent activity. The intensity of patenting is directly measured, while productivity is estimated which can lead to potential bias, but the significant finding using raw patents lends substantial credibility to our interpretation of the results from the productivity regression.

5.3 Technology or Pricing?

So far, we have studied the effect of spillovers on revenue total factor productivity which combines the influence of technology and demand factors. Without any further analysis, it is impossible to know

[22]The distribution of patents by firm is highly right-skewed and we take the logarithm to avoid our estimated being unduly affected by a few outliers and, because a number of firms have no patterns, we add unity in order to avoid taking the logarithm of zero.

[23]We use OLS estimations in this table. The large number of zeros for the dependent variable makes this regression unsuitable for weighting because a value of 0 will not change with any heteroskedasticity weighting.

[24]The presence of many firms with no patents prevents us from running a standard semi-logarithmic estimation. However, we retain the heuristic interpretation of percentage changes because we are interested in typical average effects—for a firm with no patents, the percentage change interpretation is, of course, meaningless.

whether the positive KNOWTECH and COMPTECH effects found in column (4) of Table 2 reflect changes in domestic firms' technology or prices. To see this, let us express revenue total factor productivity as:

$$\text{TFPR}_{it} \equiv P_{it}\text{TFPQ}_{\text{it}} = \mu_{\text{it}} \times \text{MC}_{\text{it}} \times \text{TFPQ}_{\text{it}}\,, \tag{21}$$

where P_{it} refers to the firms output price and TFPQ_{it} is physical productivity. Firm specific prices are available to deflate nominal output and isolate t physical productivity (TFPQ_{it}). However, in the absence of firm prices, the use of sectoral price indexes to deflate introduces biases if there are within industry deviations between firm specific prices and the sectoral index. One can write the firm price as the product of marginal cost and a markup μ_{it} as is done in equation (21). Turning to percentage changes (denoted by Δ) and re-arranging yields the following relationship between efficiency gains and changes in TFPR, markups, and marginal costs: $\Delta\text{TFPR}_{it} - \Delta\mu_{it} = \Delta\text{TFPQ}_{\text{it}} + \Delta\text{MC}_{\text{it}}$. Reordering, we obtain a formula for the change in physical productivity:

$$\Delta\text{TFPQ}_{it} \;=\; \Delta\text{TFPR}_{\text{it}} - \Delta\mu_{\text{it}} - \Delta\text{MC}_{\text{it}}\,. \tag{22}$$

If one models firm-level marginal cost and mark-ups, one can derive an estimate of changes in productivity. In order to provide an estimate of mark-ups, we follow the influential approach of De Loecker and Warzynski (2012). Hall (1986) noted that under imperfect competition, input growth is associated with disproportional output growth (as measured by the relevant markup). Based on this insight, De Loecker and Warzynski (2012) re-arrange the first order condition of the firm cost minimization problem with respect to the flexible input $\mathcal{J}$ to derive the firm markup (defined as price over marginal cost) according to this expression:

$$\mu_{it} \equiv \frac{P_{it}}{MC_{it}} = \underbrace{\frac{\partial \mathcal{F}_{it}(\cdot)}{\partial \mathcal{J}_{it}}\frac{\mathcal{J}_{it}}{\mathcal{F}_{it}(\cdot)}}_{OutputElasticity} \Big/ \underbrace{\frac{P_{it}^{\mathcal{J}_{it}}\mathcal{J}_{it}}{P_{it}\text{y}_{it}}}_{ExpenditureShare} \tag{23}$$

where P_{it} refers to the firm output price, MC_{it} stands for the firm marginal cost, $\mathcal{F}_{it}(\cdot)$ is the production function, $\mathcal{J}_{it}$ refers to the firm inputs, and $P_{it}\text{y}_{it}$ is the nominal value added.

Following De Loecker and Warzynski (2012), we consider labor as a flexible input $\mathcal{J}$. Labor's expenditure share is straightforward to compute from the data as the ratio of the labor cost to value added. The output elasticity with respect to labor is given by the elasticity obtained from the estimation of the production function β_ℓ.[25] We find empirical estimates of the median markup close to 1.5 on average.

Marin and Voigtlander (2014) express marginal cost as a function of physical productivity and input prices and show that, in their Chilean database, input prices do not change with export activity in which

[25] Notice we estimated the production function at the two-digit industry level and β_{L} is constant across two-digit industries.

case differences between revenue productivity and markups will be explained by increases in physical productivity. They further express marginal cost as an increasing function of input prices minus $\mathrm{TFPQ}_{\mathrm{it}}$. In the absence of product prices, we provide a more indirect estimate of the effect of FDI on TFPQ, which we label the effect on "implied TFPQ" by applying equation (22), the estimated mark-up, and an estimated marginal cost described next. We have data for the total material and wage cost and normalizing this with revenue, provides us with a measure of average cost. We regress this measure on our foreign ownership variable using, in particular, firm fixed effects. Regressions with firm fixed effects are identified from changes in the variables over the sample period and, while the change over time in average cost over revenue is not literally the same as marginal cost, any firm-level constants average out, and we take the estimated coefficients as estimates of the effect of the regressors on marginal cost.[26] Armed with estimates of the effect of spillovers on revenue TFP, mark-up, and marginal cost, we can provide estimates of (implied) TFPQ.

We consider the effect of each measure of FDI in turn, starting with knowledge spillovers. Column (1) in Table 4 shows our main specification with the dependent variable $\log \mathrm{TFPR}$.[27] A unit (one standard deviation) increase in KNOWTECH increases domestic firms' revenue productivity by about 0.3 percent. Column (2) uses the log of firm markup as the dependent variable and shows that a unit increase in KNOWTECH leads to an 0.1 percent increase in domestic firms' markup. Column (3) uses the average cost of domestic firms as the dependent variable and shows that a unit increase in KNOWTECH leads to a drop in marginal average cost of 0.118 percent. Column (4) shows our main object of interest, namely the effect on physical TFP. The effect of a one standard deviation change in technology weighted knowledge spillovers is 0.3 percent, which for this variable is exactly the value of the impact on revenue TFP (although less precisely estimated). Because this variable is less precisely estimated and subject to further imprecision emanating from the calculated mark-up and marginal cost, we display effects on TFPR in the majority of our tables, but the evidence points to, at the very least, a substantial part of the estimated impact is due to changes in physical productivity.

Foreign presence in the same four-digit sector increase competition as captured by COMP. The effect on TFPR of a one standard deviation change in COMP is 0.509 percent. Columns (2), (3), and (4) show that this effect can be decomposed into a 0.377 percent decrease in mark-ups, a 0.126 percent change in cost, and a 0.251 percent change in TFPQ. Because the technology weighted four-digit measure is

[26]Technically, we are using $\Delta\frac{\mathrm{cost}}{\mathrm{output}}$, minus its firm specific average. Alternatively, we use $\frac{\Delta \mathrm{cost}}{\Delta \mathrm{output}}$ calculated at the firm-level and estimate how this more direct measure of marginal cost changes with FDI. We do not obtain significant coefficients which indicates that the effect is small, but likely also that the estimate of marginal cost is noisy. We believe that in total our results point to a significant share of the effect on revenue productivity coming from physical productivity.

[27]From this point forward, we drop the variable KNOW that we have shown is not significant when the technology-weighted measure KNOWTECH is included.

included the interpretation is that of an increase in competition with no knowledge spillover while the impact of COMPTECH then is the effect of knowledge spillovers in the four-digit industry when product competition is controlled for. The predicted effect of a one standard deviation change in COMPTECH is a 0.617 percent increase in revenue productivity, and columns (2), (3), and (4) show that this effect can be decomposed into a 0.422 percent increase in mark-ups, an 0.111 percent decrease in cost, and a 0.305 percent increase in TFPQ.

To sum up, increases in revenue productivity induced by higher foreign ownership in technologically related sectors outside the same four-digit sector are mainly driven by changes in physical productivity. The effects of foreign presence in the same four digit sector has even larger effect on revenue productivity but roughly half of this is driven by changes in mark-ups and marginal costs.

5.4 Vertical Spillovers

The positive effect on firm revenue productivity, that we have identified through the variable KNOWTECH, is conceivably capturing an effect of customer-supplier relationships. We can examine this issue following the approach of the vertical spillover literature, albeit at our finer four-digit level. In order to construct vertical backward measures at the four-digit industry level, we rely on the U.S. I-O matrix coefficients for the year 2007.[28] We follow Javorcik (2004) and define the backward spillover measure as a weighted sum of the foreign presence in industries that are being supplied by sector $s4$. It is intended to capture the extent of potential contacts between domestic suppliers and multinational customers.[29] For data reasons, many studies are based on the I-O at the two-digit industry level and capture inter-**two**-digit industry spillover effects. We calculate backward spillover measures at the four-digit industry level, by differentiate between multinational customers operating within the same two-digit industry (but outside the four-digit industry) and those foreign affiliates operating in different four-digit sectors outside the same two-digit. We separately consider linkages within and outside the two-digit sector because the knowledge spillover measures used so far are defined from data within the two-digit sector and potentially could be upward biased due to left-out trade effect taking place with the two-digit industry. Further, the backward spillover effect identified in the literature are defined via FDI outside the two-digit sector and it is of interest to examine if we find similar effects—if results agree, it lends credence to both ours and previous results.

[28] Consistent I-O tables across European countries are only available at the two-digit industry classification. We assume that the production structure in the United States is similar to that in our sample of advanced European economies and match the six-digit U.S. I-O table to the four-digit NACE Rev. 2 classification used in our study.

[29] One can define $\text{FORWARD}_{s4,t}$ as a measure of foreign presence in upstream industries supplying to sector $s4$, where instead of weights $\alpha_{s4\tilde{s}4}$ we would use the weights $\sigma_{s4\tilde{s}4}$ computed as the share of inputs purchased by industry $s4$ from industry $\tilde{s}4$ in total inputs sourced by sector $s4$. The variable $\text{FORWARD}_{s4,t}$ would then capture contacts between foreign-owned suppliers and domestic customers. This measure did not came out significant in our sample, as is the case in previous work (see Table C.7).

$$\text{BACKWARD\_WITHIN2}_{s4,t} = \sum_{\substack{\tilde{s}4 \neq s4 \\ \text{if } s4, \tilde{s}4 \in s2}} \alpha_{s4\tilde{s}4} \text{HORIZONTAL}_{\tilde{s}4,t} \,,$$

and

$$\text{BACKWARD\_OUT2}_{s4,t} = \sum_{\tilde{s}4 \notin s2(s4)} \alpha_{s4\tilde{s}4} \text{HORIZONTAL}_{\tilde{s}4,t} \,,$$

where $\alpha_{s4\tilde{s}4}$ refers to proportion of sector $s4$ output supplied to sector $\tilde{s}4$ and the α coefficients are obtained from the U.S. I-O matrix.

We further control for a measure, KNOWTECH_OUT2, of closeness in technology space with firms outside the given domestic firms two-digit sector where the measure is otherwise defined in the same manner as KNOWTECH.[30] Further we control for demand and industry concentration using previously defined measures. These measures are not a focus out this paper and their inclusion serve to hedge against omitted variable bias.[31] We present the results in Table 5, which proceeds by including sequentially the measures of backward spillovers estimated at the four-digit industry level.

Table 5 about here

Column (1) shows our baseline results. Column (2) shows that including a backward measures which capture the extent of foreign presence in upstream four-digit sectors (within the same two-digit sector) does not change the size of the coefficient on the variable KNOWTECH and is statistically insignificant. This implies that the knowledge spillovers are not spuriously significant due omitted indicators of customer/supplier relationships. Column (3) includes a variable which proxies for the extent of foreign owners in other two-digit upstream manufacturing sectors. This measure, which is similar to the backward measure suggested in Javorcik (2004) and related studies in the literature, takes a positive and statistically significant coefficient. Column (5) considers the role of technology transfers from other two-digit industries in the manufacturing sector. As it can be seen, in this case, the backward measure based on I-O coefficients and the backward measure based on technology proximity are both positive and statistically significant capturing different effects on domestic firm productivity. Finally, column (6) controls for xxx compromising the robustness of the backward measure based on supplier relationships although confirming the knowledge technology closeness importance. Overall we can conclude that the knowledge

[30] We define $\text{KNOWTECH\_OUT2}_{s4,t} = \sum_{\tilde{s}4 \neq s2(s4)} \text{WTECH}_{s4,\tilde{s_4},t} \times \text{COMP}_{\tilde{s}4,t}$.

[31] We follow Javorcik (2004) and compute demand as $\text{DEMAND}_{s4,t} = \sum_{\tilde{s_4}} a_{s4,\tilde{s_4}} \times Y_{\tilde{s_4},t}$ where $a_{s4,\tilde{s_4}}$ is the I-O matrix coefficient indicating that in order to produce one unit of good $\tilde{s_4}$, $a_{s4,\tilde{s_4}}$ units of good s_{s4} are needed. $Y_{\tilde{s_4},t}$ stands for industry $\tilde{s_4}$ output deflated by an industry-specific deflator. The Herfindahl index of industry concentration is computed as: $H_{s4,t} = \sum_{i=1}^{N} ms_{i,t}^2$ where ms_i refers to the market share of firm i is four-digit industry $s4$.

measures are robust.

5.4.1 FDI Spillovers and Firm Entry and Exit

The results shown so far use an unbalanced panel of firms over the period 1999-2008, where firms can enter and exit the sample at different points in time. Bearing in mind that our dataset has some limitation for the correct identification of new entrants and exiters we explore the impact of the different spillover variables on sectoral entry and exit rates. To start, we need to differentiate "true" entry and exit from entry and exit to and from the sample (i.e., a firm may have been operating for a while without having been included in the panel before the firm at point in time enters the sample because coverage improved or because further information became available—firms may also drop from the sample because they stop reporting even if they still operate.). To identify "true" entrants into the sample, we use the year of incorporation of the firm and classified as new entrants firms that were incorporated in the corresponding current year. We classify as exiters firms that are no longer in the panel from one year onwards as well as firms that are declared bankrupt, dissolved because of bankruptcy, dissolved because of liquidation and in liquidation. While we acknowledge the limitations of these definitions, they are likely adequate for our purpose. The results in columns (1) and (2) of Table 6 are very informative on this issue. Column (1) shows that entry is lower in those sectors with high foreign activity in technologically close sectors (regardless of whether foreign investment is mainly driven by product competitors or non-product competitors). On the contrary, product-based measures of spillovers (COMP and KNOW) have a positive effect on entry. Column (2) shows that only "competition" spillover variables have an impact on the exit rate and that while product market competition reduces exit, competition on the technology space significantly increases the rate of exit. In fact, column (3) shows that the product market competition from direct or indirect competitors increases net entry while competition from foreign firms in technologically close sectors reduces net entry.

Table 6 about here

Furthermore, we find that the productivity of entrants is on average lower than the productivity of exiters, which is in turn lower than the productivity of stayers (these mean differences are statistically significant). Therefore, the extensive margin could have implications for our main results.XXXX numbers should be in table..

5.4.2 FDI Spillovers in a stable sample of firms

Table 7 replicates our main findings in terms of productivity and patents using a permanent sample of firms.[32]

Table 7 about here

The results are similar, although the estimated coefficients are larger for the permanent firms (the standardized coefficients can be compared to those of Table 2 because the standard errors of the variables in the permanent sample area similar to those of the larger sample used previously). This indicates that the results we have found so far are not driven by entry and exit of firms; in fact, the results are stronger in the permanent sample.

5.5 Timing of Impacts

It is possible that spillover effects take time to materialize in which case the results will be more economically significant at longer horizons—level regressions with firm-fixed effects are equivalent to estimating growth over the full sample. Table 8 repeats our main specification estimated in one-, two-, three- and four-year (non-overlapping) differences.[33]

Table 8 about here

In Table 8, the coefficients are not standardized because we are interested in comparing across the different differing frequencies (columns) and this would be hard to do with standardized coefficients. Further, the coefficients are simpler to interpret in economics terms and easier to compare with existing results.

Comparing the columns, we observe that knowledge spillovers become significant only after two years.[34] The coefficient to KNOWTECH after four years (see column (4)) is 5.2 which agrees very closely with the estimate in column (4) of Table 2—that estimate of 0.299 was normalized by the standard deviation of the residual and we "undo" this by dividing by the standard error given, we get a virtually

[32]The permanent sample of firms includes firms that we observe continuously from 1999 to 2008.

[33]Notice the difference specifications do not include firm fixed effects which "washes out" with differencing. We use non-overlapping data in order to avoid the mechanically induced autocorrelation that results from using overlapping data—an autocorrelation that typically cannot be well corrected for. Because we sue non-overlapping differences, the number of observations decline with the length of differencing.

[34]We chose to regress long differences rather than running regressions with many lags. While the later allows more precise statement about timing for given coefficient, for higher lags this benefit is to weighed against the many lag coefficients themselves being imprecisely estimated. On net, we prefer the long-difference specification which is also simpler to communicate with more compact sets of results.

identical coefficient of 4.98. The adjustment pattern indicates that firms need time to make use of knowledge spillovers which also agree with the finding that firms start patenting more when FDI increase in firms close in technology space. Competition effects are felt immediately as witnessed by the coefficient to COMP which are significant for one-year differencing and do not increase in size with differencing over several years. The COMPTECH measure, which is interpreted as capturing knowledge spillovers within the two-digit sector when competition is controlled for, has an effect that increases with the time span, as found for KNOWTECH, but this variable is significant already at the one year frequency.

Haskel, Pereira and Slaughter (2007) estimates spillovers from FDI in the UK—a developed country comparable to the countries in our sample—and they use differences of different length, comparable to our results. They find that spillover take time to materialize and although their results are not directly comparable with ours because the do not separate between competition and knowledge effects, nor between closeness in product or technology space (although the do consider FDI in either the same sector or the same geographical are), we can compare magnitudes. Our KNOWTECH measure increses proportionally with fo on average, so so our coefficient in column (4) of 5.2 percent can be compared to the effect found by Haskel, Pereira and Slaughter (2007) for five-year difference of 6.3 percent. Because of the differences in the specifications, we will not compare more extensively to their results, but we find it re-assuring that magnitudes agree.

6 Conclusion

TO BE ADDED

References

Acemoglu, Daron, Simon Johnson, and Todd Mitton, "Determinants of Vertical Integration: Financial Development and Contracting Costs," *Journal of Finance*, 06 2009, *64* (3), 1251–1290.

Ackerberg, Daniel A., Kevin Caves, and Garth Frazer, "Structural Identification of Production Functions," 2008. UCLA mimeo.

— , — , and — , "Identification Properties of Recent Production Function Estimators," *Econometrica*, 2015, *83* (6), 2411–2451.

Aghion, Philippe and Peter Howitt, "A Model of Growth through Creative Destruction," *Econometrica*, March 1992, *60* (2), 323–51.

— , Nicholas Bloom, Richard Blundell, Rachel Griffith, and Peter Howitt, "Competition and Innovation: An Inverted-U Relationship," *The Quarterly Journal of Economics*, 2005, *120* (2), 701–728.

Aitken, Brian J. and Ann E. Harrison, "Do Domestic Firms Benefit from Direct Foreign Investment?," *American Economic Review*, 1999, *89* (3), 605–618.

Barrios, Salvador, Holger Gorg, and Eric Strobl, "Spillovers through backward linkages from multinationals: Measurement matters!," *European Economic Review*, August 2011, *55* (6), 862–875.

Bloom, Nicholas, M. Schankerman, and John van Reenen, "Identifying Technology Spillovers and Product Market Rivalry," *Econometrica*, 2013, *4* (81), 1347–1393.

— , Mirko Draca, and John Van Reenen, "Trade Induced Technical Change? The Impact of Chinese Imports on Innovation, IT and Productivity," *Review of Economic Studies*, 2016, *83* (1), 87–117.

Gopinath, Gita, Sebnem Kalemli-Ozcan, Loukas Karabarbounis, and Carolina Villegas-Sanchez, "Capital Allocation and Productivity in South Europe," Working Papers 728, Federal Reserve Bank of Minneapolis 2015.

Hall, Robert E., "Market Structure and Macroeconomic Fluctuations," *Brookings Papers on Economic Activity*, 1986, *17* (2), 285–338.

Haskel, Jonathan E., Sonia C. Pereira, and Matthew J. Slaughter, "Does Inward Foreign Direct Investment Boost the Productivity of Domestic Firms?," *The Review of Economics and Statistics*, 2007, *89* (3), 482–496.

Helpman, Elhanan, "Trade, FDI, and the Organization of Firms," *Journal of Economic Literature*, September 2006, *44* (3), 589–630.

Jaffe, Adam B, "Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits, and Market Value," *American Economic Review*, December 1986, *76* (5), 984–1001.

—, "Demand and Supply Influences in R&D Intensity and Productivity Growth," *The Review of Economics and Statistics*, August 1988, *70* (3), 431–37.

Javorcik, Beata Smarzynska, "Does Foreign Direct Investment Increase the Productivity of Domestic Firms? In Search of Spillovers through Backward Linkages," *American Economic Review*, 2004, *94* (3), 605–627.

Kalemli-Ozcan, Sebnem, Bent Sorensen, Carolina Villegas-Sanchez, Vadym Volosovych, and Sevcan Yesiltas, "How to Construct Nationally Representative Firm Level data from the ORBIS Global Database," NBER Working Papers 21558, National Bureau of Economic Research, Inc September 2015.

Keller, Wolfgang, "International Technology Diffusion," *Journal of Economic Literature*, September 2004, *42* (3), 752–782.

— and Stephen R. Yeaple, "Multinational Enterprises, International Trade, and Productivity Growth: Firm-Level Evidence from the United States," *The Review of Economics and Statistics*, 2009, *91* (4), 821–831.

Klapper, Leora, Luc Laeven, and Raghuram Rajan, "Entry regulation as a barrier to entrepreneurship," *Journal of Financial Economics*, December 2006, *82* (3), 591–629.

Levinsohn, James and Amil Petrin, "Estimating Production Functions Using Inputs to Control for Unobservables," *Review of Economic Studies*, 2003, *70* (2), 317–342.

Loecker, Jan De and Frederic Warzynski, "Markups and Firm-Level Export Status," *American Economic Review*, 2012, *102* (6), 2437–2471.

Marin, Alvaro F. Garcia and Nico Voigtlander, "Exporting and Plant-Level Efficiency Gains: It's in the Measure," *working paper*, 2014.

Olley, G. Steven and Ariel Pakes, "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 1996, *64* (6), 1263–97.

Petrin, Amil, Jerome Reiter, and Kirk White, "The Impact of Plant-level Resource Reallocations and Technical Progress on U.S. Macroeconomic Growth," *Review of Economic Dynamics*, 2011, *14* (1), 3–26.

Rajan, Raghuram G and Luigi Zingales, "Financial Dependence and Growth," *American Economic Review*, June 1998, *88* (3), 559–86.

Wooldridge, Jeffrey M., "On Estimating Firm-Level Production Functions Using Proxy Variables to Control for Unobservables," *Economics Letters*, 2009, *104* (3), 112–114.

Table 1: Are There Positive Spillover Effects from Foreign Ownership?

DEPENDENT VARIABLE: log FIRM REVENUE TFP
SAMPLE: DOMESTIC FIRMS

	(1)	(2)	(3)
HORIZONTAL$_{s2,t-1}$	-0.084	0.515***	
	(0.104)	(0.100)	
KNOW$_{s4,t-1}$			0.155*
			(0.092)
COMP$_{s4,t-1}$			-0.115**
			(0.053)
Observations	323,730	323,730	323,730
Firm FE	✓	✓	✓
Country-Year FE	✓	✓	✓
Sec2-Year FE		✓	✓
Cluster	cs2y	cs2y	cs4y
s.d.(HORIZONTAL)	0.13	0.13	
s.d.(KNOW)			0.11
s.d.(COMP			0.16

Notes: The sample includes domestic firms, i.e., firms that have no foreign participation over the years of analysis. The table reports standardized coefficients and standard errors in parenthesis. The dependent variable is log revenue firm level productivity at time t (log $\text{TFPR}_{\text{i,t}}$). HORIZONTAL$_{s2,t-1}$ stands for horizontal spillovers and it is proxied by share of foreign output in total two-digit sectoral output. KNOW$_{s4,t-1}$ represents "knowledge spillovers" and is calculated as the share of foreign output within the same two-digit sector excluding foreign output produced within the same four-digit sector. COMP$_{s4,t-1}$ reflects "competition spillovers" and is proxied by the share of foreign output within the same four-digit sector (see section 4.1 for a full description of the regressors). All right hand side variables are lagged one period. Standard errors are clustered at the country-sector two digit-year level in columns (1) and (2) and at the country-sector four digit-year level in column (3). Results are obtained by GLS estimation using as weights the square root of the firm mean squared predicted residuals. *** denotes 1% significance; ** denotes 5% significance; * denotes 10% significance.

Table 2: Technology vs Competition Spillovers

Dependent Variable: log Firm Revenue TFP
Sample: Domestic Firms

	(1)	(2)	(3)	(4)
$\text{KNOW}_{s4,t-1}$	0.155*	-0.099	0.127	
	(0.092)	(0.114)	(0.119)	
$\text{COMP}_{s4,t-1}$	-0.115**	-0.115**	-0.508***	-0.503***
	(0.053)	(0.053)	(0.072)	(0.071)
$\text{KNOWTECH}_{s4,t-1}$		0.298***	0.244**	0.299***
		(0.079)	(0.081)	(0.066)
$\text{COMPTECH}_{s4,t-1}$			0.644***	0.619***
			(0.079)	(0.075)
Observations	323,730	323,730	323,730	323,730
Firm FE	✓	✓	✓	✓
Country-Year FE	✓	✓	✓	✓
Sec2-Year FE	✓	✓	✓	✓
Cluster	cs4y	cs4y	cs4y	cs4y
s.d.(KNOW)	0.11	0.11	0.11	
s.d.(COMP)	0.16	0.16	0.16	0.16
s.d.(KNOWTECH)		0.15	0.15	0.15
s.d.(COMPTECH)			0.06	0.06

Notes: The sample includes domestic firms, i.e., firms that have no foreign participation over the years of analysis. The table reports standardized coefficients and standard errors in parenthesis. The dependent variable is log revenue firm level productivity at time t (log $\text{TFPR}_{i,t}$). $\text{KNOW}_{s4,t-1}$ represents "knowledge spillovers" and is calculated as the share of foreign output within the same two-digit sector excluding foreign output produced within the same four-digit sector. $\text{COMP}_{s4,t-1}$ reflects "competition spillovers" and is proxied by the share of foreign output within the same four-digit sector. $\text{KNOWTECH}_{s4,t-1}$ refers to "technology weighted knowledge spillovers" and $\text{COMPTECH}_{s4,t-1}$ refers to "technology weighted competition spillovers" (see section 4.1 for a full description of the regressors). All right hand side variables are lagged one period. Standard errors are clustered at the country-sector four digit-year level. Results are obtained by GLS estimation using as weights the square root of the firm mean squared predicted residuals. *** denotes 1% significance; ** denotes 5% significance; * denotes 10% significance.

Table 3: Direct Evidence on Technology Spillovers

DEPENDENT VARIABLE: $\log(Patents + 1)$
SAMPLE: DOMESTIC FIRMS

	(1)	(2)	(3)	(4)
KNOW$_{s4,t-1}$	0.146**	0.025	0.034	
	(0.061)	(0.079)	(0.079)	
COMP$_{s4,t-1}$	-0.045	-0.045	-0.061	-0.060
	(0.035)	(0.035)	(0.045)	(0.045)
KNOWTECH$_{s4,t-1}$		0.138**	0.136**	0.153**
		(0.059)	(0.059)	(0.046)
COMPTECH$_{s4,t-1}$			0.025	0.019
			(0.048)	(0.048)
Observations	323,730	323,730	323,730	323,730
Firm FE	✓	✓	✓	✓
Country-Year FE	✓	✓	✓	✓
Sec2-Year FE	✓	✓	✓	✓
Cluster	cs4y	cs4y	cs4y	cs4y
s.d.(KNOW)	0.11	0.11	0.11	
s.d.(COMP)	0.16	0.16	0.16	0.16
s.d.(KNOWTECH)		0.15	0.15	0.15
s.d.(COMPTECH)			0.06	0.06

Notes: The sample includes domestic firms, i.e., firms that have no foreign participation over the years of analysis. The table reports standardized coefficients and standard errors in parenthesis. The dependent variable is log of the number of granted patents to firm i at time t ($\log(Patents+1)$). KNOW$_{s4,t-1}$ represents "knowledge spillovers" and is calculated as the share of foreign output within the same two-digit sector excluding foreign output produced within the same four-digit sector. COMP$_{s4,t-1}$ reflects "competition spillovers" and is proxied by the share of foreign output within the same four-digit sector. KNOWTECH$_{s4,t-1}$ refers to "technology weighted knowledge spillovers" and COMPTECH$_{s4,t-1}$ refers to "technology weighted competition spillovers" (see section 4.1 for a full description of the regressors). All right hand side variables are lagged one period. Standard errors are clustered at the country-sector four digit-year level. Results are obtained by OLS. *** denotes 1% significance; ** denotes 5% significance; * denotes 10% significance.

Table 4: Revenue TFP and Markups

Sample: Domestic Firms

Dependent Variable:	log(TFPR) (1)	log(μ) (2)	MC (3)	$ImpliedTFPQ$ (4)
KNOWTECH$_{s4,t-1}$	0.296***	0.118*	-0.118***	0.296**
	(0.059)	(0.059)	(0.030)	(0.187)
COMP$_{s4,t-1}$	-0.503***	-0.377***	0.126***	-0.251
	(0.079)	(0.063)	(0.016)	(0.285)
COMPTECH$_{s4,t-1}$	0.617***	0.422***	-0.110***	0.305***
	(0.078)	(0.065)	(0.026)	(0.078)
Observations	323730	323730	323730	
Firm FE	✓	✓	✓	
Country-Year FE	✓	✓	✓	
Sec2-Year FE	✓	✓	✓	
Cluster	cs4y	cs4y	cs4y	
s.d.(KNOWTECH)	0.15	0.15	0.15	
s.d.(COMP)	0.16	0.16	0.16	
s.d.(COMPTECH)	0.06	0.06	0.06	

Notes: The sample includes domestic firms, i.e., firms that have no foreign participation over the years of analysis. The table reports standardized coefficients and standard errors in parenthesis. In column (1), the dependent variable is log revenue firm level productivity at time t (log $\mathrm{TFPR}_{\mathrm{i,t}}$). In column (2) the dependent variable is log firm markup (log $\mu_{i,t}$). In column (3) the dependent variable is the marginal cost MC, that is computed as the sum of the cost of employment and the expenditure on materials over total output. KNOWTECH$_{s4,t-1}$ refers to "technology weighted knowledge spillovers" and COMPTECH$_{s4,t-1}$ refers to "technology weighted competition spillovers". COMP$_{s4,t-1}$ reflects "competition spillovers" and is proxied by the share of foreign output within the same four-digit sector (see section 4.1 for a full description of the regressors). All right hand side variables are lagged one period. Standard errors are clustered at the country-sector four digit-year level in column (3). Results are obtained by GLS estimation using as weights the square root of the firm mean squared predicted residuals. *** denotes 1% significance; ** denotes 5% significance; * denotes 10% significance.

Table 5: Technology Spillovers and Vertical Linkages

Dependent Variable: log Firm Revenue TFP
Sample: Domestic Firms

	(1)	(2)	(3)	(4)	(5)
KNOWTECH$_{s4,t-1}$	0.296***	0.281***	0.370***	0.370***	0.326***
	(0.059)	(0.074)	(0.074)	(0.074)	(0.074)
COMP$_{s4,t-1}$	-0.503***	-0.503***	-0.503***	-0.519***	-0.471***
	(0.079)	(0.079)	(0.079)	(0.079)	(0.079)
COMPTECH$_{s4,t-1}$	0.617***	0.644***	0.677***	0.703***	0.794***
	(0.078)	(0.078)	(0.078)	(0.078)	(0.078)
BACKWARD_WITHIN2$_{s4,t-1}$		0.042	0.035	-0.007	0.035
		(0.049)	(0.049)	(0.057)	(0.057)
BACKWARD_OUT2$_{s4,t-1}$		0.146**	0.120**	0.051	0.051
		(0.051)	(0.051)	(0.051)	(0.051)
KNOWTECH_OUT2$_{s4,t-1}$			0.989***	1.054***	1.126***
			(0.245)	(0.245)	(0.245)
log DEMAND$_{s4,t-1}$				0.942***	0.837***
				(0.209)	(0.209)
log HERFIN$_{s4,t-1}$					-0.582***
					(0.116)
Observations	323,730	322,523	322,523	322,523	322,523
Firm FE	✓	✓	✓	✓	✓
Country-Year FE	✓	✓	✓	✓	✓
Sec2-Year FE	✓	✓	✓	✓	✓
Cluster	cs4y	cs4y	cs4y	cs4y	cs4y
s.d.(KNOWTECH)	0.15	0.15	0.15	0.15	0.15
s.d.(COMP)	0.16	0.16	0.16	0.16	0.16
s.d.(COMPTECH)	0.06	0.07	0.07	0.07	0.07
s.d.(BACKWARD_WITHIN2)		0.07	0.07	0.07	0.07
s.d.(BACKWARD_OUT2)		0.09	0.09	0.09	0.09
s.d.(KNOWTECH_OUT2)			0.07	0.07	0.07
s.d.(log DEMAND)				1.05	1.05
s.d.(log HERFIN)				1.16	1.16

Notes: The sample includes domestic firms, i.e., firms that have no foreign participation over the years of analysis. The table reports standardized coefficients and standard errors in parenthesis. The dependent variable is log revenue firm level productivity at time t (log $\text{TFPR}_{\text{i,t}}$). KNOWTECH$_{s4,t-1}$ refers to "technology weighted knowledge spillovers" and COMPTECH$_{s4,t-1}$ refers to "technology weighted competition spillovers". COMP$_{s4,t-1}$ reflects "competition spillovers" and is proxied by the share of foreign output within the same four-digit sector. BACKWARD_WITHIN2$_{s4,t-1}$ refers to backward vertical spillovers within the same two-digit sector. BACKWARD_OUT2$_{s4,t-1}$ refers to backward vertical spillovers outside the two-digit sector. KNOWTECH_OUT2$_{s4,t-1}$ proxies for technology weighted backward spillovers outside the two-digit sector. log DEMAND$_{s4,t-1}$ is the log of four-digit sector demand. log HERFIN$_{s4,t-1}$ is the log of four-digit Herfindahl index. All right hand side variables are lagged one period. Standard errors are clustered at the country-sector four digit-year level. Results are obtained by GLS estimation using as weights the square root of the firm mean squared predicted residuals. *** denotes 1% significance; ** denotes 5% significance; * denotes 10% significance.

Table 6: FDI Spillovers and Selection

Sector Level Results

	(1)	(2)	(3)
Dependent Variable:	Entry	Exit	Net Entry
KNOWTECH$_{s4,t-1}$	-0.433***	0.038	-0.569***
	(0.057)	(0.057)	(0.058)
COMP$_{s4,t-1}$	0.251***	-0.183***	0.593***
	(0.023)	(0.023)	(0.046)
COMPTECH$_{s4,t-1}$	-0.057*	0.625***	-1.696***
	(0.032)	(0.049)	(0.170)
KNOW$_{s4,t-1}$	0.890***	0.058	0.552***
	(0.058)	(0.058)	(0.049)
Observations.	5100	5100	5100
Country-Sec4	✓	✓	✓
Country-Year FE	✓	✓	✓
Sec2 FE	✓	✓	✓
Cluster	cs4y	cs4y	cs4y

Notes: The entry and exit rates are based on the sample of domestic firms, i.e., firms that have no foreign participation over the years of analysis. The dependent variable in column (1) is the entry rate in each four-digit sector in year t. The dependent variable in column (2) is the exit rate in each four-digit sector in each year t. The dependent variable in column (3) is the difference between the entry and the exit rate at the four-digit sector level in each year t KNOW$_{s4,t-1}$ represents "knowledge spillovers" and is calculated as the share of foreign output within the same two-digit sector excluding foreign output produced within the same four-digit sector. COMP$_{s4,t-1}$ reflects "competition spillovers" and is proxied by the share of foreign output within the same four-digit sector. KNOWTECH$_{s4,t-1}$ refers to "technology weighted knowledge spillovers" and COMPTECH$_{s4,t-1}$ refers to "technology weighted competition spillovers" (see section 4.1 for a full description of the regressors). All right hand side variables are lagged one period. Standard errors are clustered at the country-sector four digit-year level. Results are obtained by GLS estimation using as weights the square root of the sector mean squared predicted residuals. *** denotes 1% significance; ** denotes 5% significance; * denotes 10% significance.

Table 7: Permanent Sample of firms

SAMPLE: DOMESTIC FIRMS

Dependent Variable:	log(TFPR)	$\log(1 + Patents)$
$\text{COMP}_{s4,t-1}$	-0.904***	-0.081
	(0.129)	(0.081)
$\text{KNOWTECH}_{s4,t-1}$	0.451**	0.180**
	(0.150)	0.075
$\text{COMPTECH}_{s4,t-1}$	1.304***	0.070
	(0.146)	(0.070)
Observations	101875	101875
Firm FE	✓	✓
Country-Year FE	✓	✓
Sec2-Year FE	✓	✓
Cluster	cs4y	cs4y
s.d.(KNOWTECH)	0.15	0.15
s.d.(COMP)	0.16	0.16
s.d.(COMPTECH)	0.06	0.06

Notes: Sample of domestic firms that have no foreign participation over the years of analysis. The table reports standardized coefficients and standard errors in parenthesis. The dependent variable in column (1) is log revenue firm level productivity at time t (log $\text{TFPR}_{\text{i,t}}$) and results are estimated by GLS using as weights the square root of the firm mean squared predicted residuals. The dependent variable in column (2) is the log of the number of patents plus one and is estimated by OLS. $\text{KNOWTECH}_{s4,t-1}$ refers to "technology weighted knowledge spillovers" and $\text{COMPTECH}_{s4,t-1}$ refers to "technology weighted competition spillovers". $\text{COMP}_{s4,t-1}$ reflects "competition spillovers" and is proxied by the share of foreign output within the same four-digit sector. All right hand side variables are lagged one period. Standard errors are clustered at the country-sector four digit-year level. *** denotes 1% significance; ** denotes 5% significance; * denotes 10% significance.

Table 8: FDI Spillovers: Speed of Adjustment

Sample: Domestic Firms

	(1)	(2)	(3)
Dependent Variable:	$\Delta^{j=1}\log\mathrm{TFPR}$	$\Delta^{j=2}\log\mathrm{TFPR}$	$\Delta^{j=4}\log\mathrm{TFPR}$
$\Delta^j\mathrm{KNOWTECH}_{s4}$	0.3	2.2***	5.2***
	(0.6)	(0.3)	(0.4)
$\Delta^j\mathrm{COMP}_{s4}$	-3.1***	-3.2***	-2.8***
	(0.7)	(0.5)	(0.4)
$\Delta^j\mathrm{COMPTECH}_{s4}$	7.4***	10.8***	13.4***
	(1.9)	(1.4)	(1.1)
Observations	323,730	172,525	72,652
Firm FE			
Country-Year FE	✓	✓	✓
Sec2-Year FE	✓	✓	✓
Cluster	cs4y	cs4y	cs4y

Notes: The sample includes domestic firms, i.e., firms that have no foreign participation over the years of analysis. The table reports non-standardized coefficients and standard errors in parenthesis multiplied by 100. The dependent variable is the corresponding non-overlapping year difference in log revenue firm level productivity at time t ($\Delta^j\log\mathrm{TFPR}_{\mathrm{i,t}}$) where j indicates, first, second and forth differences. $\mathrm{KNOWTECH}_{s4,t-1}$ refers to "technology weighted knowledge spillovers" and $\mathrm{COMPTECH}_{s4,t-1}$ refers to "technology weighted competition spillovers". $\mathrm{COMP}_{s4,t-1}$ reflects "competition spillovers" and is proxied by the share of foreign output within the same four-digit sector (see section 4.1 for a full description of the regressors). Standard errors are clustered at the country-sector four digit-year level. Results are obtained by GLS estimation using as weights the square root of the firm mean squared predicted residuals. *** denotes 1% significance; ** denotes 5% significance; * denotes 10% significance.

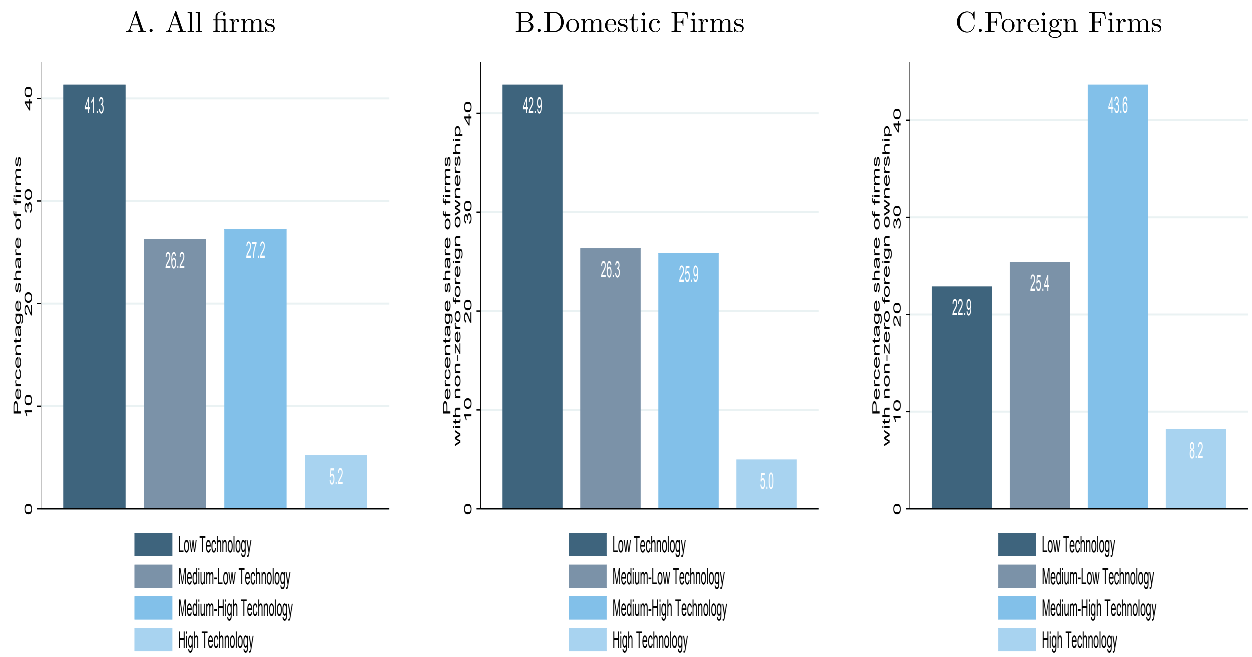
Table 9: Predicted Effects of a 10 Percentage Points Increase in FDI

SAMPLE: DOMESTIC FIRMS

	Across the Board	FDI in Highly Connected Sectors	FDI in Less Connected Sectors
Predicted Effect	0.42	0.58	0.25

Notes: Column (1) displays the predicted average increase in TFPR in manufacturing as a results of a 10 percentage point increase in FDI, equally distributed across sectors. Column (2) displays the predicted average increase in TFPR in manufacturing as a result of a 10 percentage point increase in FDI, distributed as a 20 percentage point increase in FDI in half of sectors, chosen as those that are closer to other sectors on average and 0 percentage points in the more isolated sectors. "Closeness to other sectors on average" means that ..xxyyzz. Column (3) shows the predicted impact of the same amount of FDI, but focused in the sectors that are below the median in terms of being close to others in technology space on average. (MORE. ALSO COLUMNS WITH EXTREMELY CONNECTED OR ISOLATED)

Figure 1: Distribution of firms and foreign firms in all firms sample, split by technological intensity of sectors



Notes: The shares of firms in individual bars are computed relative to the total number of firms in sample and are thus additive.

Figure 2: Foreign Ownership stakes by technological intensity of sectors

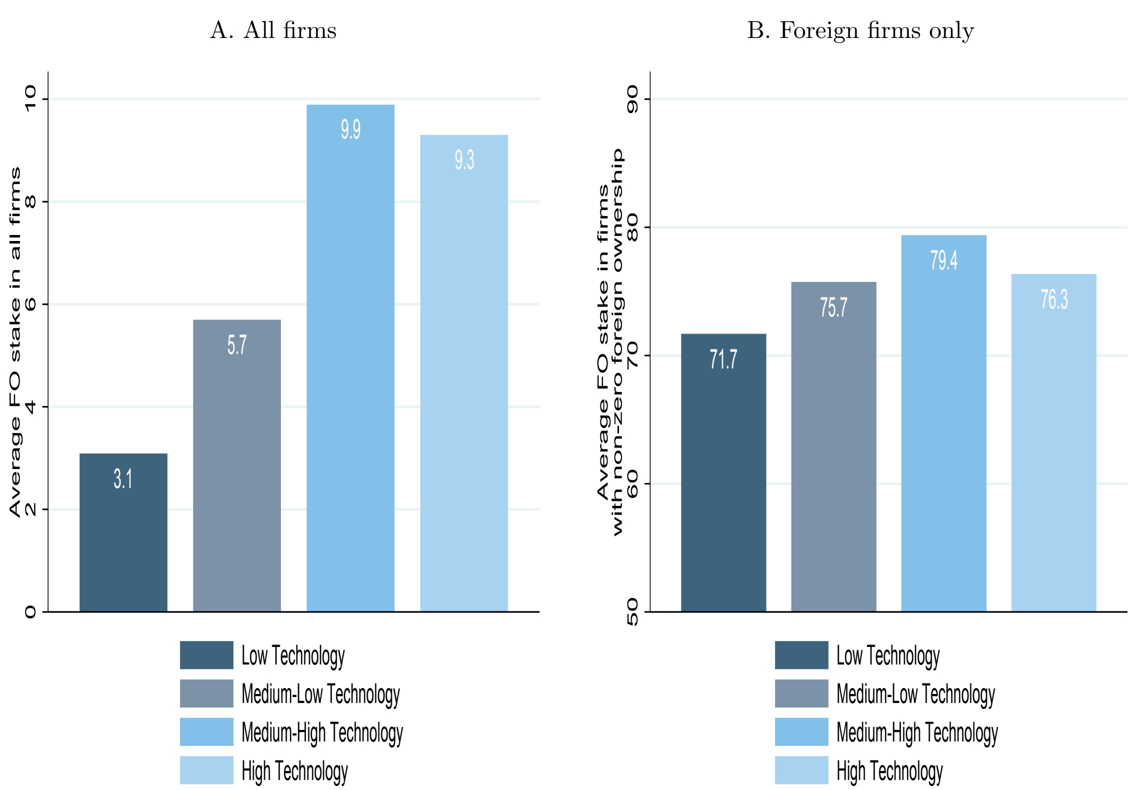


Figure 3: Distribution of R&D weights by technology intensity of sectors

A. All values

B. Truncated at 99 percentile

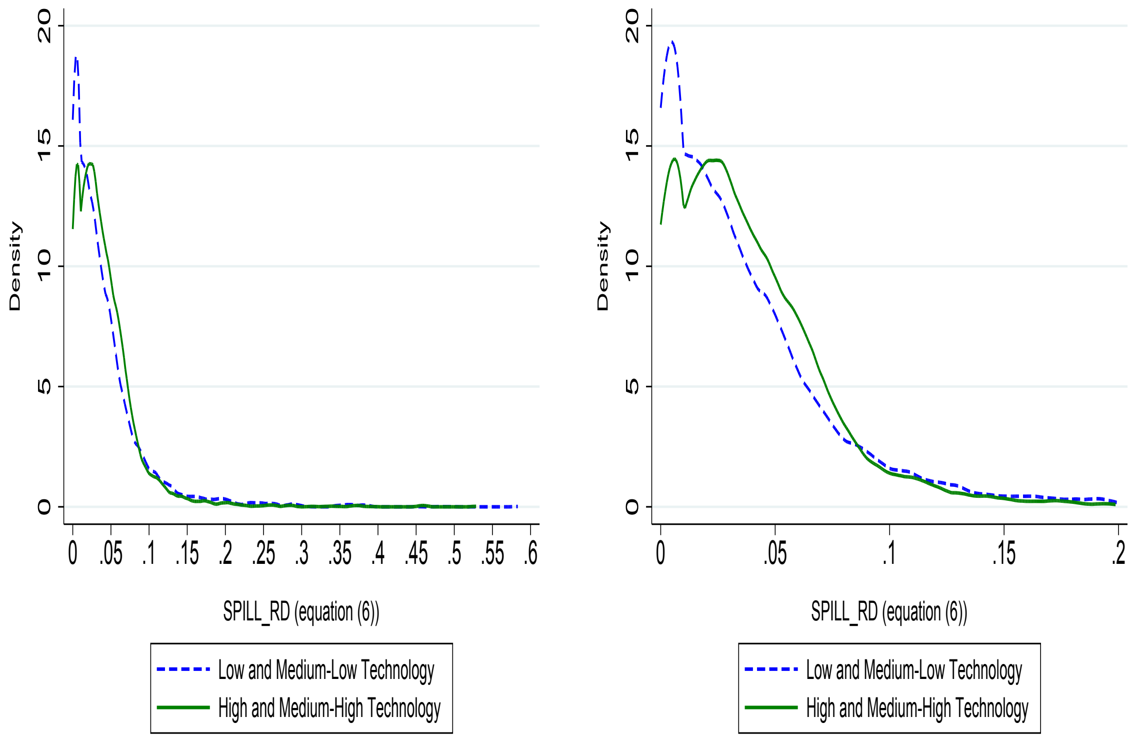


Figure 4: Spillover measures by technology sector in 2007

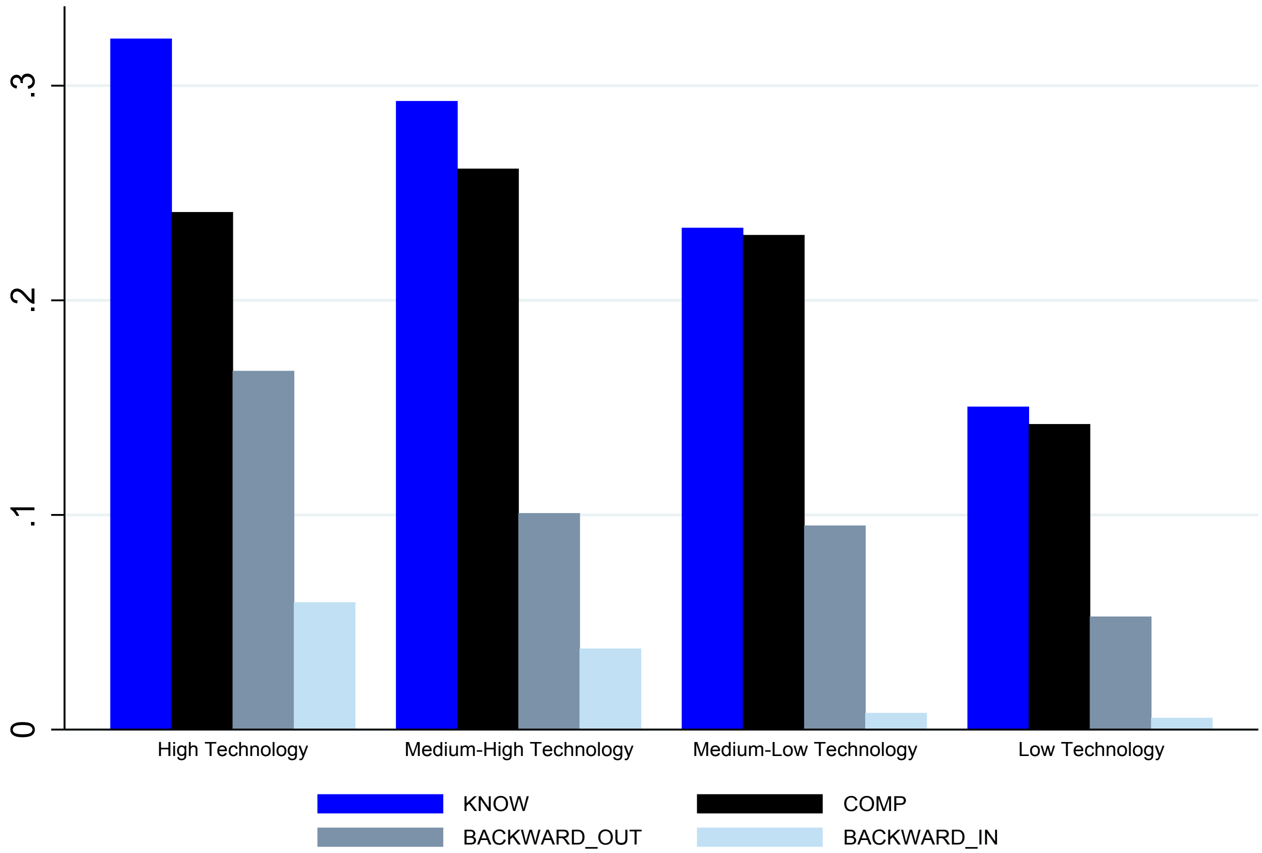


Figure 5: Correlation KNOW and KNOWTECH. Four-digit sector in year 2007.

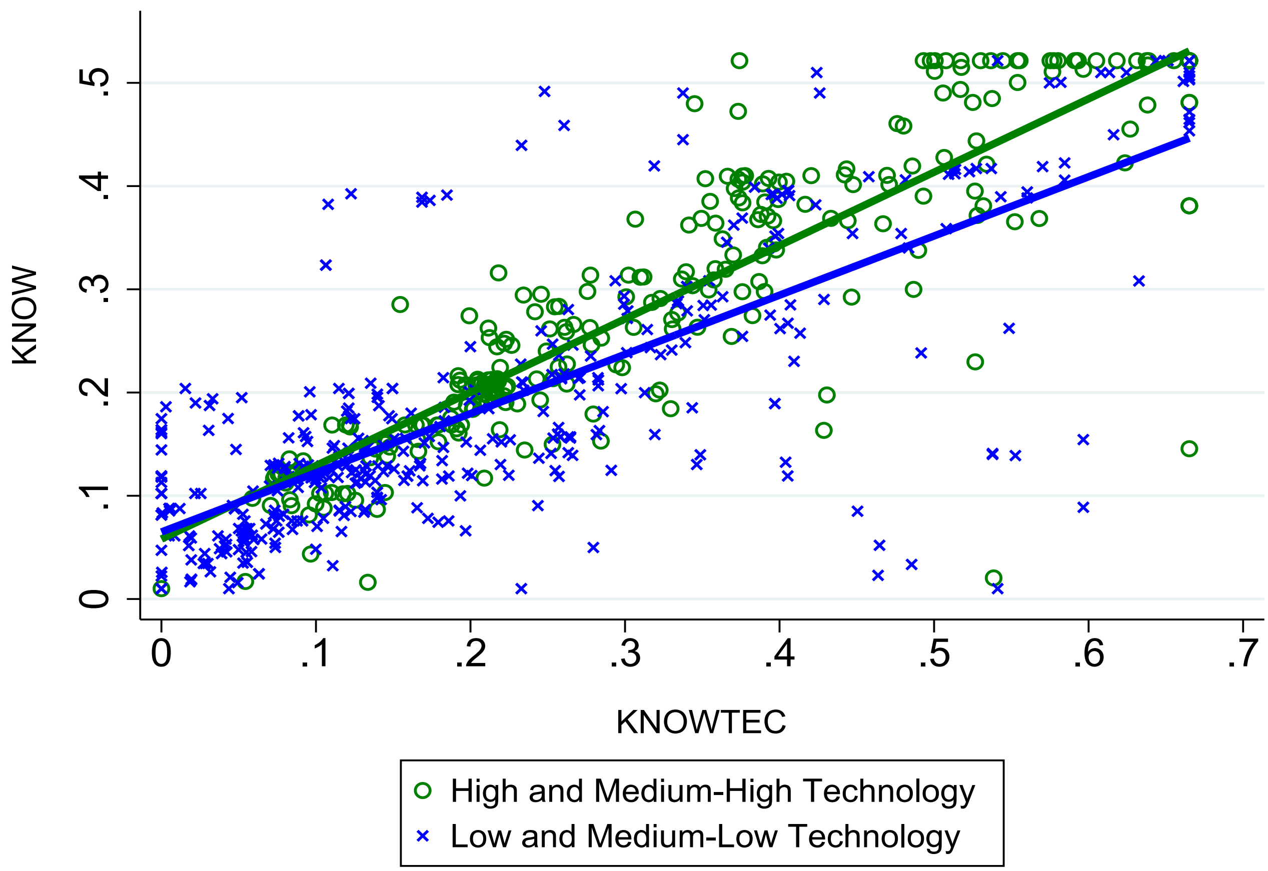
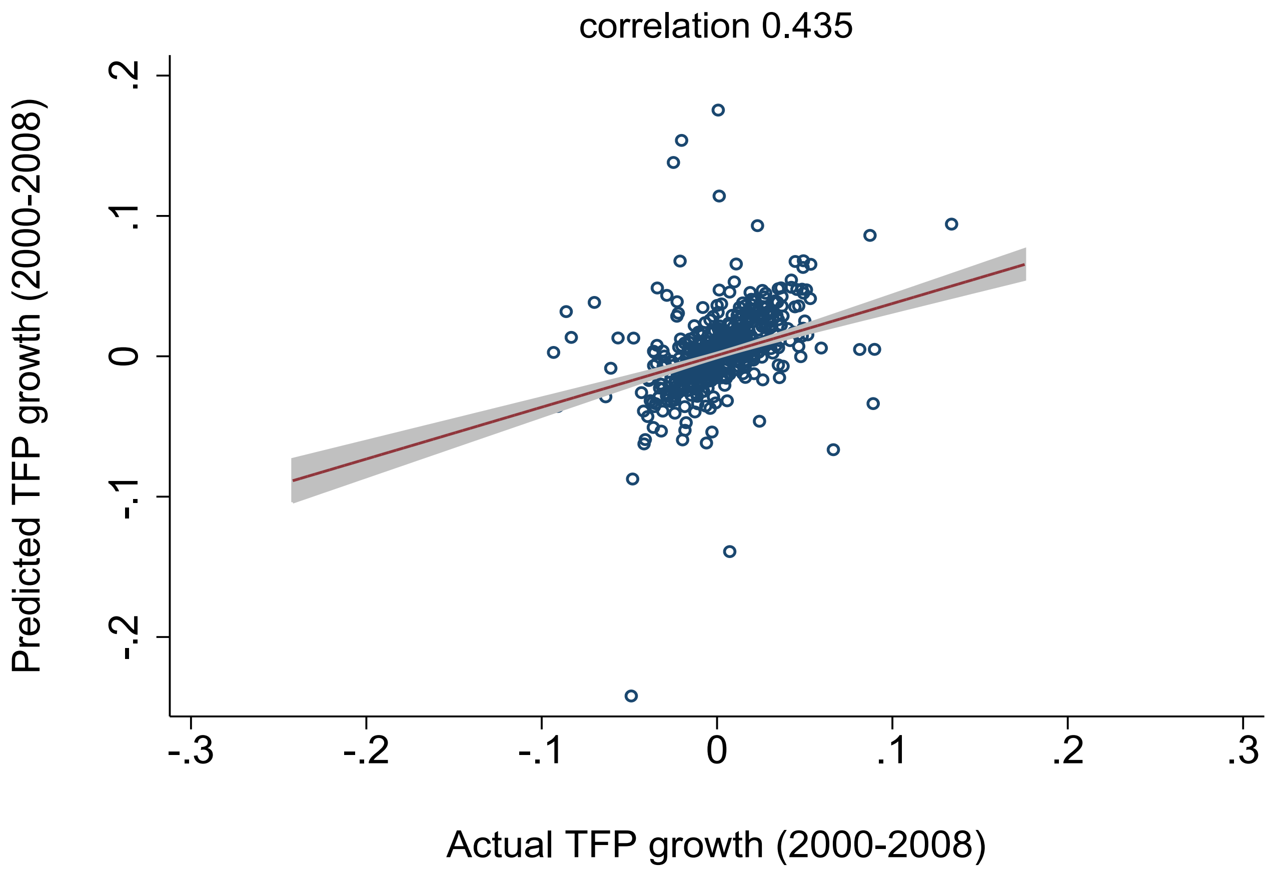


Figure 6: Average TFP growth and predicted TFP growth between 2000 and 2008



Appendix

A Data

In Kalemli-Ozcan, Sorensen, Villegas-Sanchez, Volosovych and Yesiltas (2015) we provided the comprehensive comparison of our data relying on the ORBIS database to the best available national sources. Here we briefly touch upon some statistics relevant for the countries in our study and refer the reader to Kalemli-Ozcan et al. (2015) for more details.

Output (Turnover). Table A.1 shows how much of the official gross output data from Eurostat's *Structural Business Statistics* (SBS) database is covered by the firms in our data for the total economy (Panel A) and manufacturing sector (Panel B). Each cell is the ratio of value of total output produced by our firms relative to value of total output produced as in the official data. BvD provides firm-level information on gross output (represented by Turnover) for all sectors of a given European country between 1999-2008, however Eurostat SBS data provides information on gross output (Turnover) for a subset of sectors. So, for a given country-year, ratios are computed by taking the ratio of aggregated gross output values where aggregated gross output is computed by cumulating gross output over these sectors for which the gross-output related variable is available in both data sets. Missing ratios still appear in some country-year due to missing Eurostat data. As shown in Panel A our data can account more than 50 percent of the aggregate output in Finland and Italy, close to 60 percent in Belgium, and above 70 percent in other countries. Our data cover even larger portion of the manufacturing output as shown in Panel B. In our comparison study we also demonstrate that we do a good job covering firms reporting the data necessary for estimation of the TFP, such as employment or wage bill, tangible fixed assets, and cost of materials. In addition, our data is representative when we count employment and compare that to the numbers from Eurostat's SBS data as well as in terms of number of firms in the economy relative to Eurostat and the sector-level CompNet database that is constructed by the ECB from similar sources as BvD data used by us.[35]

Important for our study is the representativeness in terms of firm sizes, within the manufacturing sector. Table A.2 reports the size distribution in the manufacturing sector by gross output accounted for by firms belonging to three size categories in a random year 2006. Row entries denote the fraction of total economic activity accounted for by firms belonging to each size class. Here we also match well the official statistics in terms of size distribution of economic activity undertaken in manufacturing sector.

Foreign Ownership. We validate our foreign ownership data by comparing it to the country-level *Activities of Foreign Affiliates* (AFA) database by the OECD, covering "the affiliates under foreign control" (available at http://stats.oecd.org/Index.aspx?DataSetCode=AFA_IN3). The data is provided,

[35]The comparison by number of businesses is less informative because many firms in Eurostat have zero employment (self-employed) while neither our database nor CompNet includes self-employed. Even with this caveat though, our data captures most of the firms in the total economy in terms of number of firms.

largely, at one-digit level of ISIC rev. 3 or 4 classification, but allows to estimate the share of foreign output in manufacturing. This data is not directly comparable to our firm-level statistics, where our definition of foreign investors refers to the direct owners in other countries and we observe all, including very small, stakes, often by multiple foreign entities. According to the OECD, in their data the ultimate owner of the domestic affiliate should be considered, but in some countries it is the immediate controlling entity. In addition, the notion of foreign affiliate in OECD-AFA is based on the concept of controlling interest, which varies across countries. In most countries, the controlling interest is based on majority ownership (50%) while others also consider minority control (between 10% and 50%). Moreover, some countries include indirectly owned foreign affiliates in addition to the directly controlled ones.

For example, in Finland the OECD provides information on direct foreign owners but multiple foreign control is not considered. In France the identification of affiliates of foreign groups and their accounting variables come from two different surveys. The Italian survey focuses on foreign ultimate owners with controlling stakes. Norwegian data is subject to several methodology changes. For example, the indirectly foreign-owned establishments are covered from 2002 onwards. Until 2003, the annual survey on foreign assets and liabilities (the Financial Census) was the main source for updating the register of foreign assets and liabilities (SIFON), while the Directorate of Taxes register of shareholders has been the main source since 2004. The press and Internet etc. are also used to map new foreign-controlled enterprises in Norway, and to map the ultimate country of ownership. The data collection methodology Spain is not well documented at all and the data for Belgium is available in AFA in only a single year (due to the country reporting restrictions). Because the coverage of most in-sample countries is problematic in AFA prior to 2002 we present the resulting time series for 2002–07 and exclude Belgium. Our BvD data implies that foreign shares for Belgium are broadly in line with the aggregated statistics (though are above average), therefore, excluding this country won't mis-represent the averages.

We compute the foreign presence in total turnover based on our (BvD) data as the average of the four-digit sectoral spillover measure $\textsc{comp}_{s4,t}$ in equation (4) over all sectors and countries in our sample. We compare the time series in this measure with the average share of the turnover of the foreign affiliates in total manufacturing turnover reported at the country level as directly available from the OECD. Alternatively we compute the foreign part of turnover as the ratio of the foreign part of turnover, aggregated over firm-sector-countries, to the total turnover. That is, instead of averaging the expression $\textsc{comp}_{s4,t}$ in equation (4) we compute $\frac{\sum_{s4,c}(\sum_{i\in s4,c}\text{fo}_{i,c,t}\times\text{go}_{i,c,t})}{\sum_{s4,c}(\sum_{i\in s4}\text{go}_{i,c,t})}$, making sure that we retain the sample of the firms which remain in-sample in all years (the "permanent firms" sample). This latter measure, presumably, is less affected by the individual firms or sector outliers. In case of OECD data, we recover the monetary value of multinational turnover from the ratios reported by the OECD, express them in a single currency, add them up across countries, and then divide by the total monetary value of overall manufacturing turnover.

Figure A.1 presents the results with average shares in Panel A and shares computed from aggregated data in Panel B. Both figures imply that our data represents the foreign presence reported by government agencies to the OECD very well. Finally, In Figure A.2 we present the data behind the Panel B in the previous figure, but now for individual countries. All five figures have the same vertical scale for comparability. With notable exception of Spain, we represent the level and the trends in the OECD data well for our in-sample countries.

In Kalemli-Ozcan, Sorensen, Villegas-Sanchez, Volosovych and Yesiltas (2015) we compare the number, output, and employment of foreign affiliates covered by BvD and OECD and demonstrate that in most countries our data covers more than 50 percent of the multinational economic activity reported in the OECD and in many countries ORBIS-AMADEUS does better. Table A.3 reports those statistics for our 5 countries (Belgium is missing in OECD). As seen, ORBIS-AMADEUS has better coverage in countries such as Finland, Italy, or Norway where we observe the ratios bigger than 1.

Table A.1: Coverage Based on Gross Output (Turnover), BvD vs. Eurostat Data

	Belgium	Finland	France	Italy	Spain	Norway
A: Total Economy						
1999	0.60	0.35	0.54	0.45	0.63	0.63
2000	0.65	0.40	0.57	0.50	0.64	0.63
2001	0.63	0.44	0.65	0.50	0.67	0.77
2002		0.47	0.62	0.54	0.67	0.79
2003	0.63	0.50	0.67	0.53	0.68	0.65
2004	0.63	0.51	0.70	0.57	0.68	0.67
2005	0.63	0.53	0.69	0.57	0.69	0.59
2006	0.62	0.51	0.68	0.58	0.71	0.67
2007	0.62	0.55	0.70	0.60	0.68	0.71
2008	0.73	0.57	0.79	0.72	0.80	0.59
B: Manufacturing Sector						
1999	0.75	0.3	0.64	0.61	0.75	0.60
2000	0.8	0.34	0.76	0.66	0.77	0.60
2001	0.78	0.36	0.79	0.65	0.78	
2002		0.37	0.82	0.71	0.8	0.75
2003	0.81	0.39	0.79	0.7	0.79	0.68
2004	0.8	0.41	0.83	0.73	0.79	0.72
2005	0.8	0.41	0.82	0.77	0.78	0.69
2006	0.78	0.4	0.84	0.79	0.83	0.75
2007	0.79	0.45	0.87	0.79	0.81	0.76
2008	0.78	0.49	0.9	0.9	0.85	0.69

Notes: Table presents the ratio that are calculated based on gross output for the in-sample countries. The sample consists of firms that report data with positive values of gross output (Turnover). Panel A covers all available sectors in an economy. BvD provides firm-level information on gross output (Turnover) for all sectors of a given European country, however Eurostat SBS data provides information on gross output with the exceptions of some sectors. So, for a given country-year, total economy percentages are computed by taking the ratio of aggregated gross output values where aggregated gross output is computed by totalling gross output over these sectors for which gross-output related variable is available in both data sets. Panel B covers manufacturing sector.

Table A.2: Size Distribution by Gross Output (Turnover) in Manufacturing Sector, 2006, BvD vs. Eurostat Data

	Belgium	Finland	France	Italy	Spain	Norway
A: ORBIS-AMADEUS						
1 to 19 employees	0.05	0.08	0.05	0.12	0.13	0.11
20 to 249 employees	0.30	0.38	0.23	0.49	0.40	0.40
250 + employees	0.66	0.54	0.72	0.40	0.47	0.49
B: EUROSTAT (SBS)						
0 to 19 employees	0.08	0.06	0.09	0.20	0.14	0.13
20 to 249 employees	0.27	0.21	0.27	0.41	0.38	0.36
250 + employees	0.65	0.74	0.64	0.39	0.49	0.51

Notes: Table presents the share of gross output accounted for by firms belonging in three size categories in the year 2006. The sample consists of firms that report data with positive values of gross output (Turnover). Panel A reports the measures from our data based on ORBIS-AMADEUS and in Panel B are same numbers from Eurostat's SBS database. Row entries denote the fraction of total economic activity accounted for by firms belonging to each size class. Each column is a different country.

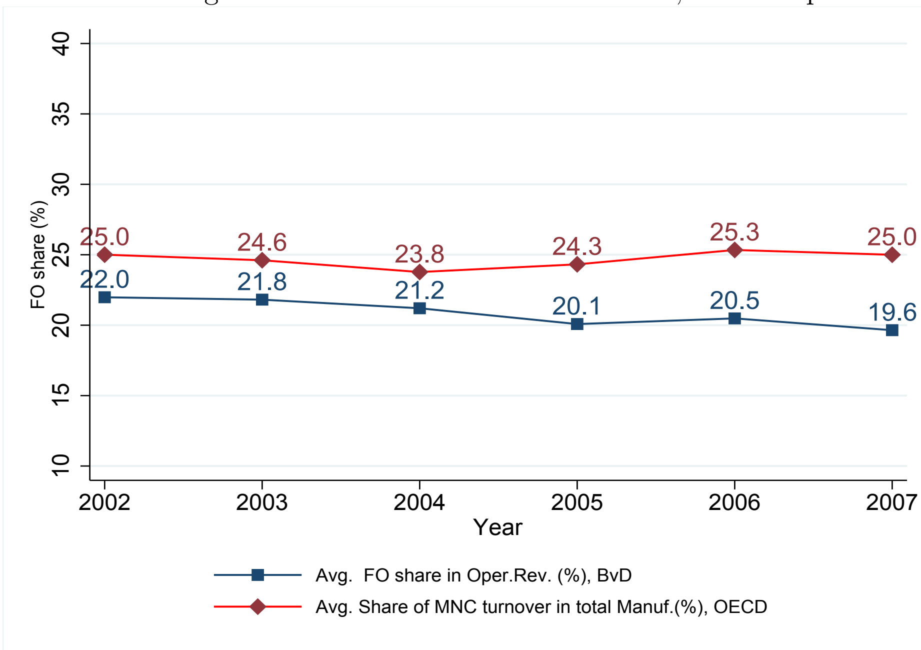
Table A.3: Foreign Affiliate Coverage in Manufacturing, BvD vs. OECD Data

	Finland		France		Italy		Spain		Norway	
	FO10	FO50	FO10	FO50	FO10	FO50	FO10	FO50	FO10	FO50
Panel A: Number of Firms										
1999	0.61	0.48	0.60	0.55	.	.	1.45	1.16	0.37	0.29
2000	0.65	0.51	0.64	0.59	.	.	1.44	1.16	0.55	0.44
2001	0.63	0.50	0.66	0.60	0.92	0.70	1.31	1.07	0.52	0.41
2002	0.83	0.68	0.53	0.49	1.08	0.81	1.03	0.87	0.52	0.41
2003	0.78	0.64	0.51	0.47	0.98	0.74	1.06	0.90	0.52	0.40
2004	0.97	0.79	0.48	0.44	1.23	0.91	0.95	0.83	0.56	0.42
2005	0.95	0.77	0.46	0.42	1.23	0.91	0.87	0.76	0.57	0.43
2006	0.87	0.70	0.49	0.44	1.39	1.02	0.93	0.81	0.56	0.42
2007	0.82	0.67	0.53	0.47	1.63	1.18	0.89	0.77	0.71	0.51
2008	0.90	0.74	0.57	0.50	1.89	1.35	0.84	0.73	0.83	0.57
Panel B: Output										
1999	0.19	0.18	0.37	0.34	.	.	0.76	0.60	0.34	0.28
2000	0.22	0.20	0.42	0.39	.	.	0.78	0.61	0.43	0.36
2001	0.26	0.24	0.42	0.38	0.62	0.51	0.79	0.64	0.39	0.34
2002	0.35	0.34	0.40	0.36	0.66	0.54	0.66	0.60	0.32	0.27
2003	0.46	0.37	0.38	0.35	0.67	0.54	0.64	0.58	0.47	0.28
2004	2.63	1.21	0.32	0.30	0.80	0.68	0.64	0.59	1.67	0.89
2005	2.37	1.06	0.31	0.29	0.93	0.81	0.65	0.59	1.77	0.89
2006	2.22	0.99	0.38	0.34	0.85	0.73	0.71	0.66	1.54	0.77
2007	1.92	0.85	0.39	0.35	0.83	0.73	0.71	0.65	1.69	0.84
2008	1.76	0.86	0.77	0.37	1.18	0.84	0.86	0.70	1.75	1.00

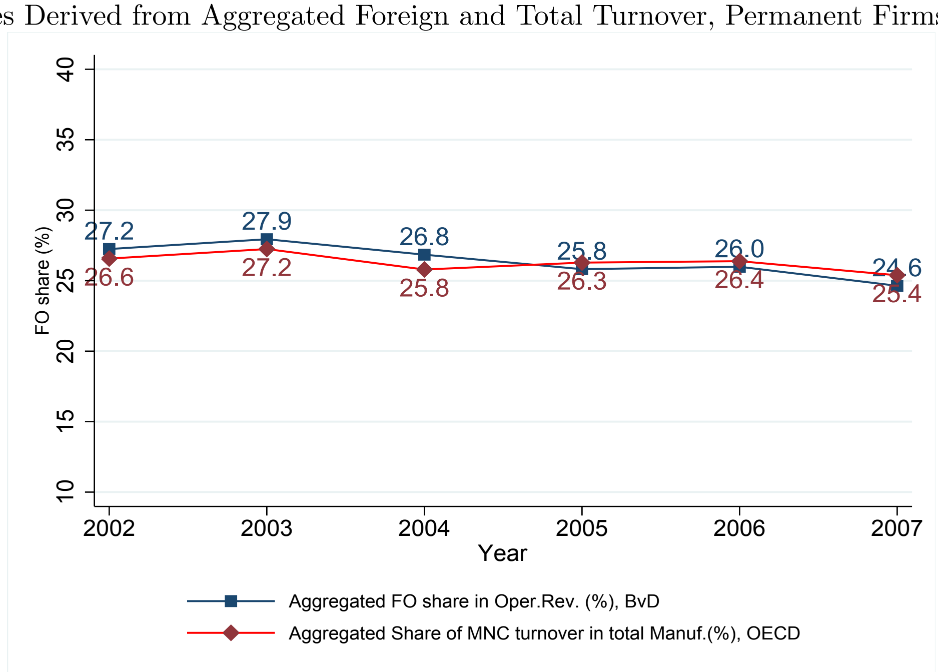
Notes: The table presents the ratio of outcome reported in BvD to that in OECD data. $FO10$ refers to companies with more than 10% foreign ownership. $FO50$ refers to companies with more than 50% foreign ownership. Panel A reports the number of firms and Panel B shows percentages in terms of output (Turnover).

Figure A.1: Foreign Shares in Turnover: BvD vs. OECD Data

A: Average Shares over Sectors and Countries, Full Sample

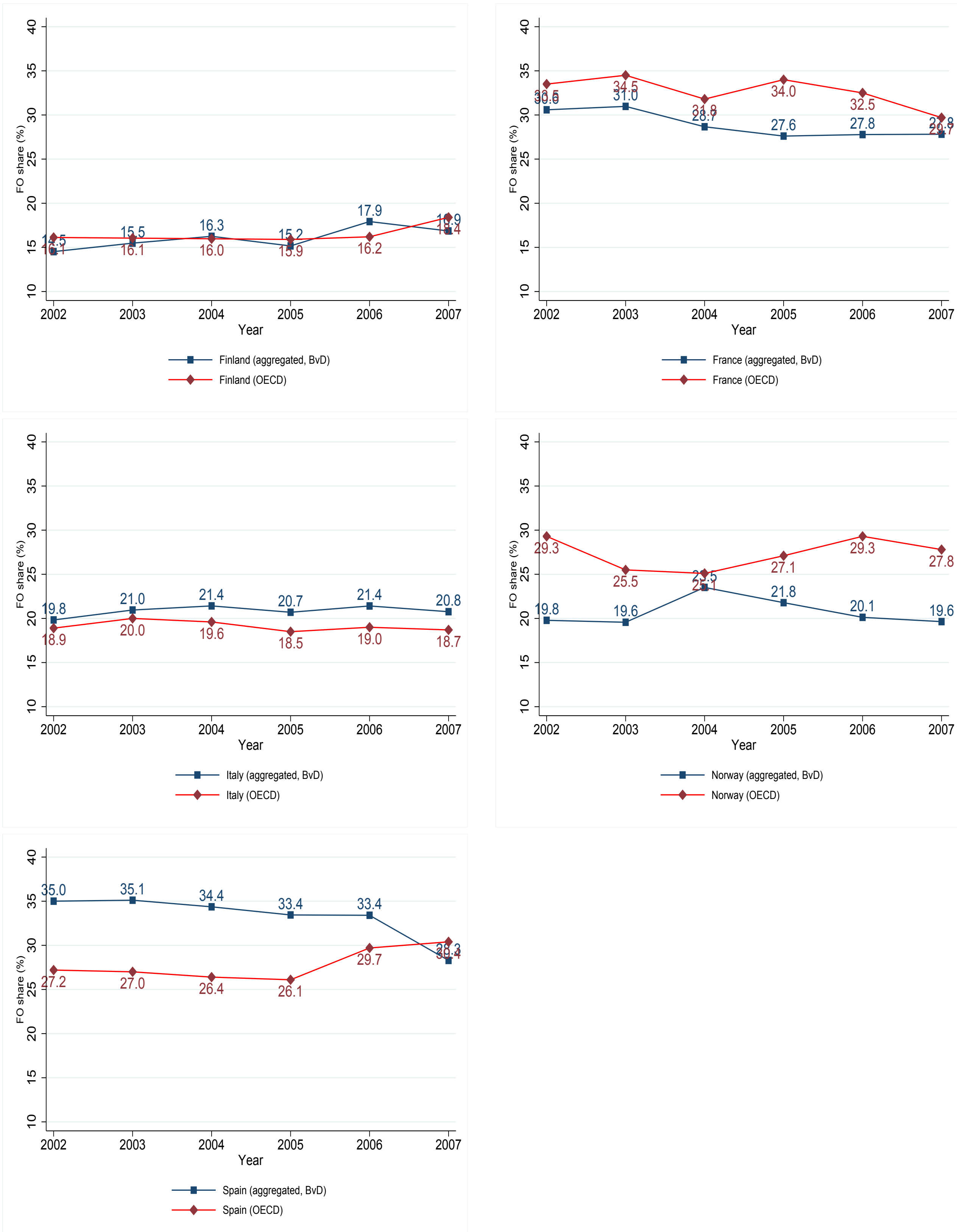


B: Shares Derived from Aggregated Foreign and Total Turnover, Permanent Firms Sample



Notes: All ratios are in percent. In Panel A, the foreign presence from the BvD data is computed as the average of the four-digit sectoral spillover measure $\textsc{comp}_{s4,t}$ in equation (4) over all sectors and countries in our sample and foreign presence from the OECD data is the average share of the country-lervel multinational turnover in total manufacturing turnover. In Panel B, all the shares from the BvD data are computed as the ratios of the aggregated foreign turnover computed as $\frac{\sum_{s4,c}(\sum_{i\in s4,c} \mathrm{fo}_{i,c,t}\times \mathrm{go}_{i,c,t})}{\sum_{s4,c}(\sum_{i\in s4} \mathrm{go}_{i,c,t})}$ over firms i, sectors $s4$ and countries c in the balanced (permanent) firm sample and foreign presence from the OECD data is the sum of the multinational turnover in manufacturing over countries divided by the total manufacturing turnover in these countries. Countries included are Finaland, France, Italy, Norway and Spain. Belgium is omitted due to lack of the OECD data.

Figure A.2: Foreign Shares in Turnover: BvD vs. OECD Data, individual countries



Notes: All ratios are in percent. All the shares from the BvD data are computed as the ratios of the aggregated foreign turnover computed as $\frac{\sum_{s4}(\sum_{i\in s4,c} \mathrm{fo}_{i,c,t}\times \mathrm{go}_{i,c,t})}{\sum_{s4}(\sum_{i\in s4} \mathrm{go}_{i,c,t})}$ over firms i and sectors $s4$ in the balanced (permanent) firm sample and foreign presence from the OECD data is the ratio of multinational turnover in manufacturing divided by the total manufacturing turnover in these countries. Belgium is omitted due to lack of the OECD data.

B Production Function Estimation

B.1 Methodology

This appendix explains the details of the firm-level productivity estimation performed using the method of Wooldridge, Levinsohn and Petrin, as suggested by Olley and Pakes (1996) and Levinsohn and Petrin (2003) and further augmented by Wooldridge (2009). Olley and Pakes (1996) (OP) and Levinsohn and Petrin (2003) (LP) propose to use proxy variables to control for unobserved productivity. The estimation in both methods is based on a two-step procedure to achieve consistency of the coefficient estimates for the inputs of the production function. Wooldridge (2009) suggests a generalized method of moments estimation of TFPR to overcome some limitations of OP and LP, including correction for simultaneous determination of inputs and productivity, no need to maintain constant returns to scale, and robustness to the Ackerberg et al. (2008) critique.[36] The following discussion is based on Wooldridge (2009), accommodated to the case of a production functions with two production inputs (see Wooldridge (2009) for a general discussion).

For firm i in time period t define:

$$y_{it} = \alpha + \beta_l l_{it} + \beta_k k_{it} + \omega_{it} + e_{it}\,, \tag{B.1}$$

where y_{it}, l_{it}, and k_{it} denote the natural logarithm of firm value added, labor (a variable input), and capital, respectively. The firm specific error can be decomposed into a term capturing firm specific productivity ω_{it} and an additional term that reflects measurement error or unexpected productivity shocks e_{it}. We are interested in estimating ω_{it}.

A key assumption of the OP and LP estimation methods is that for some function $g(.,.)$:

$$\omega_{it} = g(k_{it}, m_{it})\,, \tag{B.2}$$

where m_{it} is a proxy variable (for investment in OP, for intermediate inputs in LP). Under the assumption,

$$E(e_{it}|l_{it}, k_{it}, m_{it}) = 0 \qquad t = 1, 2, ..., T\,, \tag{B.3}$$

substituting equation ((B.2)) into equation ((B.1)), we have the following regression function:

$$\begin{aligned} E(y_{it}|l_{it}, k_{it}, m_{it}) &= \alpha + \beta_l l_{it} + \beta_k k_{it} + g(k_{it}, m_{it}) \\ &\equiv \beta_l l_{it} + h(k_{it}, m_{it})\,, \end{aligned} \tag{B.4}$$

[36] Ackerberg et al. (2008) highlight that if the variable input (labor) is chosen prior to the time when production takes place, the coefficient on variable input is not identified.

where $h(k_{it}, m_{it}) \equiv \alpha + \beta_k k_{it} + g(k_{it}, m_{it})$.

In order to identify β_l and β_k, we need some additional assumptions. First, rewrite equation ((B.3)) in a form allowing for more lags :

$$E(e_{it}|l_{it}, k_{it}, m_{it}, l_{i,t-1}, k_{i,t-1}, m_{i,t-1}, ..., l_{i1}, k_{i1}, m_{i1}) = 0 \qquad t = 1, 2, ..., T\,. \tag{B.5}$$

Second, assume productivity follows a first-order Markov process:

$$E(\omega_{it}|\omega_{i,t-1}, ..., \omega_{i1}) = E(\omega_{it}|\omega_{i,t-1}) \qquad t = 2, 3, ..., T, \tag{B.6}$$

and assume that the productivity innovation $a_{it} \equiv \omega_{it} - E(\omega_{it}|\omega_{i,t-1})$ is uncorrelated with current values of the state variable k_{it} as well as past values of the variable input l, the state k, and the proxy variables m:

$$\begin{aligned} &E(\omega_{it}|k_{it}, l_{i,t-1}, k_{i,t-1}, m_{i,t-1}, ..., l_{i1}, k_{i1}, m_{i1}) \\ &\quad = E(\omega_{it}|\omega_{i,t-1}) \equiv f[g(k_{i,t-1}, m_{i,t-1})]\,. \end{aligned} \tag{B.7}$$

Recall from equation((B.2)) that $\omega_{i,t-1} = g(k_{i,t-1}, m_{i,t-1})$.

Plugging $\omega_{i,t} = f[g(k_{i,t-1}, m_{i,t-1})] + a_{it}$ into equation ((B.1)) gives:

$$y_{it} = \alpha + \beta_l l_{it} + \beta_k k_{it} + f[g(k_{i,t-1}, m_{i,t-1})] + a_{it} + e_{it}\,. \tag{B.8}$$

Now it is possible to specify two equations which identify (β_l, β_k):

$$y_{it} = \alpha + \beta_l l_{it} + \beta_k k_{it} + g(k_{i,t}, m_{i,t}) + e_{it} \tag{B.9}$$

and

$$y_{it} = \alpha + \beta_l l_{it} + \beta_k k_{it} + f[g(k_{i,t-1}, m_{i,t-1})] + u_{it}\,, \tag{B.10}$$

where $u_{it} \equiv a_{it} + e_{it}$.

Important for the GMM estimation strategy, the available orthogonality conditions differ across these two equations. The orthogonality conditions for equation ((B.9)) are those outlined in the equation((B.5)), while the orthogonality conditions for equation ((B.10)) are

$$E(u_{it}|k_{it}, l_{i,t-1}, k_{i,t-1}, m_{i,t-1}, ..., l_{i1}, k_{i1}, m_{i1}) = 0 \qquad t = 2, ..., T\,. \tag{B.11}$$

To proceed with the estimation, we estimate these equations parametrically. In that, we follow Petrin et al. (2011) and use a third-degree polynomial approximation using first order lags of variable input as instruments.[37]

B.2 Estimation Results

Table B.1 reports summary statistics of the output elasticities estimated using the Wooldridge (2009) approach. Results are consistent across countries with no major differences except for Belgium where the number of observations is slightly lower and the coefficient on labor is on average marginally lower (0.601) and the average coefficient on capital marginally higher (0.113). Summary statistics are computed excluding sectors in which the WLP procedure delivers either missing, negative or zero coefficients. These cases are a minority and mainly correspond to sector "12. Manufacture of Tobacco products" and sector "19. Manufacture of coke and refined petroleum products" that have very few observations and a minor contribution to the overall manufacturing output.

Table B.1: Summary Statistics of the Production Function Output Elasticities

	Labor Elasticity (β_ℓ)	Capital Elasticity (β_k)
Mean	0.719	0.083
Median	0.725	0.079
Standard Deviation	0.099	0.057
Max	0.943	0.573
Min	0.133	0.004

[37]We use the Stata routine suggested in Petrin et al. (2011).

C Appendix Tables and Figures

Table C.1: Employment Characteristics Across Samples

Country	Sample	Mean	Median	SD
Total	original	61.4	26	157.1
	merged	65.1	27	172.6
Belgium	original	151.6	55	410.3
	merged	168.0	57	499.2
Spain	original	57.4	25	124.9
	merged	58.5	25	136.3
Finland	original	56.4	24	113.7
	merged	57.4	24	115.2
France	original	75.5	27	202.9
	merged	85.4	30	227.1
Italy	original	49.2	25	100.2
	merged	50.5	25	101.1
Norway	original	54.8	24	121.9
	merged	58.2	24	133.9

Table C.2: Summary Statistics

	Observations	Mean	SD
$\log VA$	519,516	14.62	1.35
$\log L$	519,516	13.72	1.30
$\log K$	519,516	13.21	1.77
$\log TFPR$	519,516	3.49	0.90
HORIZONTAL	519,516	0.21	0.13
COMP	519,516	0.17	0.16
KNOW	519,516	0.18	0.11
KNOWTECH	519,516	0.22	0.15
COMPTECH	490,193	0.05	0.07
BACKWARD_WITHIN2	517,619	0.02	0.07
BACKWARD_OUT2	519,235	0.10	0.09
KNOWTECH_OUT2	519,516	0.29	0.07

Table C.3: Summary Statistics

	Observations	Mean	SD	Min	Max	Skewness	Kurtosis
VA	363,617	6672335	14600000	30723.99	6.13E+08	8.60	171.80
L	363,617	2467715	5017241	8708.747	2.04E+08	8.37	158.07
K	363,617	2680382	6851640	3115.833	2.99E+08	8.53	172.31
M	363,617	7090686	19500000	3061.916	1.18E+09	15.10	621.47
$TFPR$	363,617	60.18	1311.84	3.487763	314897.30	184.77	36328.39
HORIZONTAL	363,617	0.21	0.13	0	0.86	1.00	3.81
COMP	363,617	0.17	0.16	0	0.71	1.36	4.48
KNOW	363,617	0.18	0.11	0.01	0.52	0.89	3.33
KNOWTECH	363,617	0.22	0.15	0	0.66	0.82	3.14
COMPTECH	343,353	0.05	0.07	0	0.36	2.58	10.53
BACKWARD_WITHIN2	362,217	0.02	0.07	0	0.96	8.38	85.23
BACKWARD_OUT2	363,446	0.10	0.09	0	1	2.74	17.09
KNOWTECH_OUT2	363,617	0.29	0.07	0.17	0.56	1.05	4.88

Notes: These are the summary stats for the non-log values. Monetary values are deflated. The sample is the one we use in regressions (where the lag for the basic spillover variables is non-missing).

Table C.4: Summary Statistics

	Observations	Mean	SD	Min	Max	Skewness	Kurtosis
VA	363,617	14.84	1.29	10.33	20.23	0.01	3.45
L	363,617	13.94	1.23	9.07	19.13	-0.13	3.86
K	363,617	13.43	1.72	8.04	19.52	-0.09	2.89
M	363,617	14.42	1.78	8.03	20.88	-0.46	3.71
$TFPR$	363,617	3.53	0.92	1.25	12.66	-0.12	3.59
HORIZONTAL	363,617	0.18	0.10	0	0.62	0.75	3.08
COMP	363,617	0.15	0.13	0	0.54	1.08	3.57
KNOW	363,617	0.16	0.09	0.01	0.42	0.69	2.93
KNOWTECH	363,617	0.19	0.12	0	0.51	0.58	2.65
COMPTECH	343,353	0.05	0.06	0	0.30	2.35	9.11
BACKWARD_WITHIN2	362,217	0.02	0.05	0	0.67	7.31	67.08
BACKWARD_OUT2	363,446	0.09	0.07	0	0.69	2.02	10.73
KNOWTECH_OUT2	363,617	0.25	0.05	0.16	0.44	0.85	4.18

Notes: These are the summary stats for the log values. Monetary values are deflated. The sample is the one we use in regressions (where the lag for the basic spillover variables is non-missing). The log of the spillovers variables is computed as $log(variable + 1)$.

Table C.5: Correlation

	log TFPR	HORIZONTAL	COMP	KNOW	KNOWTECH	COMPTECH	BACKWARD_WITHIN2	BACKWARD_OUT2	KNOWTECH_OUT2
log TFPR	1								
HORIZONTAL	0.001	1							
COMP	-0.005	0.234	1						
KNOW	-0.002	0.833	-0.142	1					
KNOWTECH	0.005	0.585	-0.099	0.679	1				
COMPTECH	0.005	0.211	0.698	-0.280	-0.169	1			
BACKWARD_WITHIN2	0.003	0.214	-0.064	0.322	0.484	-0.125	1		
BACKWARD_OUT2	0.008	-0.045	-0.039	-0.013	-0.003	-0.057	-0.068	1	
KNOWTECH_OUT2	0.000	-0.242	-0.070	-0.197	-0.254	-0.078	-0.121	0.143	1

Table C.6: Technology vs Competition Spillovers: OLS vs GLS

DEPENDENT VARIABLE: log FIRM REVENUE TFP
SAMPLE: DOMESTIC FIRMS

Estimation Method:	OLS (1)	GLS (2)	OLS (3)	GLS (4)	OLS (5)	GLS (6)	OLS (7)	GLS (8)
$\text{KNOW}_{s4,t-1}$	0.045	0.157*	-0.392	-0.101	-0.112	0.123		
	(0.224)	(0.090)	(0.303)	(0.112)	(0.314)	(0.123)		
$\text{COMP}_{s4,t-1}$	-0.251*	-0.110**	-0.251*	-0.110**	-0.738***	-0.503***	-0.738***	-0.503***
	(0.141)	(0.047)	(0.141)	(0.047)	(0.189)	(0.079)	(0.189)	(0.079)
$\text{KNOWTECH}_{s4,t-1}$			0.503**	0.296***	0.444**	0.237**	0.385**	0.296***
			(0.207)	(0.074)	(0.222)	(0.074)	(0.163)	(0.059)
$\text{COMPTECH}_{s4,t-1}$					0.773***	0.643***	0.786***	0.617***
					(0.208)	(0.078)	(0.195)	(0.078)
Observations	323,730	323,730	323,730	323,730	323,730	323,730	323,730	323,730
Firm FE	✓	✓	✓	✓	✓	✓	✓	✓
Country-Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Sec2-Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Cluster	cs4y	cs4y	cs4y	cs4y	cs4y	cs4y	cs4y	cs4y
s.d.(KNOW)	0.11	0.11	0.11	0.11	0.11	0.11		
s.d.(COMP)	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16
s.d.(KNOWTECH)			0.15	0.15	0.15	0.15	0.15	0.15
s.d.(COMPTECH)					0.06	0.06	0.06	0.06

Notes: The sample includes domestic firms, i.e., firms that have no foreign participation over the years of analysis. The table reports standardized coefficients and standard errors in parenthesis. The dependent variable is log revenue firm level productivity at time t (log $\text{TFPR}_{\text{i,t}}$). $\text{KNOW}_{s4,t-1}$ represents "knowledge spillovers" and is calculated as the share of foreign output within the same two-digit sector excluding foreign output produced within the same four-digit sector. $\text{COMP}_{s4,t-1}$ reflects "competition spillovers" and is proxied by the share of foreign output within the same four-digit sector. $\text{KNOWTECH}_{s4,t-1}$ refers to "technology weighted knowledge spillovers" and $\text{COMPTECH}_{s4,t-1}$ refers to "technology weighted competition spillovers" (see section 4.1 for a full description of the regressors). All right hand side variables are lagged one period. Standard errors are clustered at the country-sector four digit-year level. Results are obtained by OLS estimation in columns (1), (3), (5) and (7) and by GLS estimation using as weights the square root of the firm mean squared predicted residuals in columns (2), (4), (6) and (8). *** denotes 1% significance; ** denotes 5% significance; * denotes 10% significance.

Table C.7: Forward Spillovers

Dependent Variable: log Firm Revenue TFP
Sample: Domestic Firms

	(1)	(2)	(3)
KNOW$_{s4,t-1}$	2.3***		
	(0.8)		
COMP$_{s4,t-1}$	-0.5	-3.2***	-3.0***
	(0.3)	(0.5)	(0.5)
BACKWARD$_{s4,t-1}$	1.1**	1.1**	0.5
	(0.5)	(0.5)	(0.5)
FORWARD$_{s4,t-1}$	0.2	0.9	1
	(0.8)	(0.8)	(0.8)
KNOWTECH$_{s4,t-1}$		2.1***	1.9***
		(0.5)	(0.5)
COMPTECH$_{s4,t-1}$		10.1***	11.9***
		(1.2)	(1.2)
KNOWTECH_OUT2		11.6***	13.4***
		(3.3)	(3.3)
log DEMAND$_{s4,t-1}$			0.8***
			(0.2)
log HERFIN$_{s4,t-1}$			-0.5***
			(0.1)
Observations	342,832	323,730	323,730
Firm FE	✓	✓	✓
Country-Year FE	✓	✓	✓
Sec2-Year FE	✓	✓	✓
Cluster	cs4y	cs4y	cs4y

Notes: The sample includes domestic firms, i.e., firms that have no foreign participation over the years of analysis. The dependent variable is log revenue firm level productivity at time t (log $\text{TFPR}_{i,t}$). KNOWTECH$_{s4,t-1}$ refers to "technology weighted knowledge spillovers" and COMPTECH$_{s4,t-1}$ refers to "technology weighted competition spillovers". COMP$_{s4,t-1}$ reflects "competition spillovers" and is proxied by the share of foreign output within the same four-digit sector. BACKWARD$_{s4,t-1}$ refers to backward vertical spillovers. FORWARD$_{s4,t-1}$ refers to forward vertical linkages. KNOWTECH_OUT2 proxies for technology weighted backward spillovers outside the two-digit sector. log DEMAND$_{s4,t-1}$ is the log of four-digit sector demand. log HERFIN$_{s4,t-1}$ is the log of four-digit Herfindahl index. All right hand side variables are lagged one period. Standard errors are clustered at the country-sector four digit-year level. Results are obtained by GLS estimation using as weights the square root of the firm mean squared predicted residuals. *** denotes 1% significance; ** denotes 5% significance; * denotes 10% significance.

Figure C.1: Spillover Measures by Two-Digit sector in 2007

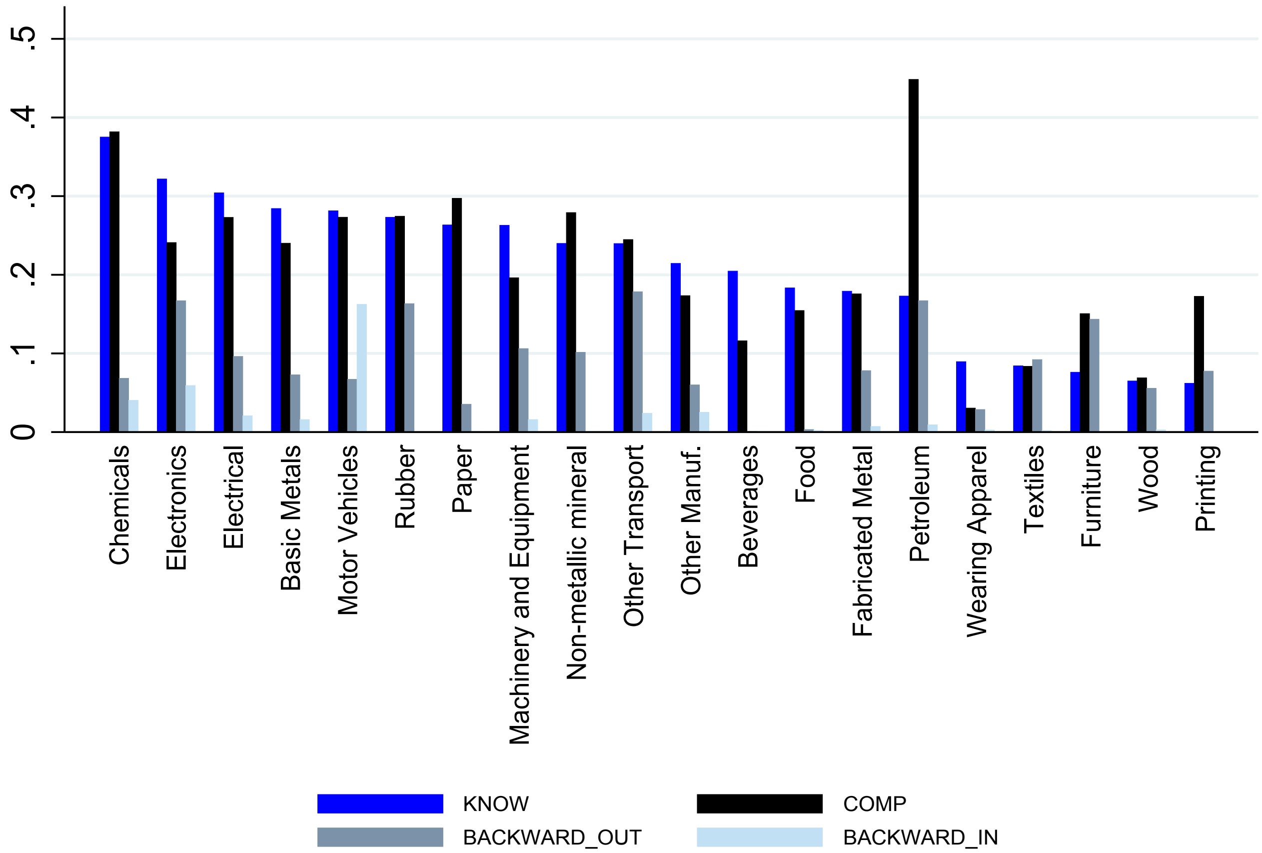


Figure C.2: Spillover Measures by Two-Digit sector in 2007 (excluding sector 19)