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

September 2016

Abstract

We show that the mixed results found in the literature that estimates horizontal technology spillovers from foreign investment on domestic firms' productivity are due to confounding the effects of positive knowledge spillovers with the negative effects of competition between foreign and domestic firms. We propose a new measure to identify knowledge spillovers on domestic firms' productivity from foreign investment in the same sector using a unique firm level data set from six advanced European countries. Our data set is nationally representative and provides us with information to link firms internationally via their subsidiaries and also to link domestic and foreign firms at a very granular level. Our "knowledge spillover" measure eliminates the direct competition effect between foreign and domestic firms by focusing on foreign affiliates operating within the same *two-digit* sector but outside the same *four-digit* sector of a given domestic firm. We find positive horizontal knowledge spillover effects only after the direct competition effect is eliminated and selection effects of foreign investment into sectors over time are controlled for via the use of sector-year dummies. We further weight the knowledge spillover measure by the "technological closeness" measure of **?** and show that our knowledge spillover measure captures a technology transfer. Our results show that horizontal knowledge spillovers coming from technological know-how can be much larger than what has been previously estimated: an XX percent increase in domestic firms' productivity as a result of being in the same horizontal sector with a foreign firm that operates with similar technology.

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

A stylized fact in the international macro literature is that foreign direct investment (FDI) promotes growth. Many theoretical studies have shown that the presence of multinational companies in an economy affect the productivity of purely domestic firms by creating "spillovers." While there is a robust positive correlation between FDI and growth in aggregate cross-country data, the literature that investigates the mechanism behind this effect using firm-level data delivers mixed results. The evidence in favor of multinationals creating "vertical spillovers," via customer-supplier relationships in different sectors linked via an input-output matrix is strong. However, the evidence supporting positive "horizontal spillovers" that is, spillovers among similar firms operating in the same sector is mixed.[1] Such "horizontal technology/knowledge spillovers," can be broadly defined as learning about technology, mimicking of production or sales practices, or mobile workers bringing skills acquired in foreign affiliates to domestic firms. There is an inherent downward bias in the estimates of "horizontal spillovers" given the competition effects between foreign and domestic firms to capture a bigger market share. Such competition effects will put a downward pressure on output prices and/or an upward pressure on input prices, biasing estimates of measured productivity for domestic firms downward.

In the literature, the typical "horizontal spillover" regression will regress the productivity of domestic firms on the presence of multinationals in the sector that domestic firms operate in. Multinational presence is measured as the share of foreign firms' output in total industry output, where typically a two-digit industry classification has been used. There are two key reasons why such an approach will fail to disentangle "knowledge spillovers" from competition effects. First, the "horizontal spillover" proxy is characterized by mismeasurement, where within the same two-digit industry, firms do not necessarily need to be competitors for market shares calling for a more disaggregated industry classification when computing "horizontal spillovers." Even foreign and domestic firms are direct competitors in their final good production, there can still be simultaneous positive "knowledge" and negative "competition" effects derived from foreign investment within the same two-digit sector. Second, the literature so far has investigated one country at a time. Given the fact that multinationals' presence is measured at the sector level, this approach cannot control for the reasons of multinational entry, i.e., the selection effect. For example, the transportation equipment sector might grow faster than the average sector, which is why the transportation equipment sector receives foreign investment in the first place, or basic metal sector might grow slower than the average sector so firms in that sector might be available for sale at cheap prices for a global energy firm.

To deal with these concerns we propose new measures and a methodology that will account for

[1] See **?** for evidence on vertical spillovers and see **?** for a summary of mixed results on horizontal spillovers and the literature in general.

sector-selection. Our new measure will capture "knowledge spillovers" based on a new unique dataset on international firm linkages across six advanced European economies. Our data set has detailed information regarding all foreign affiliates and subsidiaries of multinationals and domestic firms operating in all these countries and crucially for our study, information on the sector of activity of domestic and foreign firms is available at the four-digit industry level. Our approach allows us to contribute to the literature in two major ways. First, we breakdown the traditional two-digit industry "horizontal technology spillover" into competition effects and "knowledge spillovers." "Knowledge spillovers" represent the share of foreign firms' output operating within the same two-digit sector as the domestic firm but outside the same four-digit sector. We expect "knowledge spillovers" to be free of competition effects. If domestic and foreign firms operate in the same two-digit sector but in different four-digit sectors then they are not direct product market competitors and the domestic firm can benefit from the technology advantage of the foreign firm. Second, instead of the standard approach of using a single country, our multi-country, multi sector firm level data set lends itself to control for differential growth rates across sectors, i.e., sector-selection, with the use of two-digit sector-year fixed effects.

Our benchmark results show that there is a positive effect of "knowledge spillovers" and an insignificant effect of competition on the domestic firms' revenue total factor productivity (TFPR). To understand these results further, we follow **?** (BSvR) who propose a method for estimating technology spillovers and competition effects from research and development (R&D) activity. Their methodology is based on the similarity of firms in the "technology space" and in the "product space." Our foreign competition measure is similar to BSvR's product market competition measure since it captures closeness in the product space. To check whether our new foreign "knowledge spillover" measure, proxies for knowledge transfer, we re-calculate our output weighted "knowledge spillover" measure by re-weighting output of foreign firms in different four-digit sectors with BSvR's measure of bilateral sector similarity in patenting and R&D expenditures. We compute these weights by aggregating the cross-firm measures constructed by BSvR for listed U.S. firms to the sectoral classification in our dataset. The advantage of using U.S. firms is that U.S. patent citations and R&D expenditures are exogenous to foreign investment in Europe. We show that these "technology-weighted knowledge" spillover measures renders our basic "knowledge spillover" variable insignificant. In addition, our competition measure turns out to be negative and significant. Realizing the possibility that there can be learning even between the direct competitors, we further weight our competition measure with the BSvR's technology weights. Our basic competition measure remains to be negative significant but now the "technology-weighted competition" variable is positive and significant. This exercise confirms the existence of simultaneous positive spillover effects and negative competition effects from foreign direct competitors operating within the same four-digit sector, where positive knowledge spillover effects from direct competitors come only from those competitors that are technologically similar to the given domestic firm.

Our results show that the economic effects of horizontal knowledge and technology spillovers from foreign firms to domestic firms operating in the same sector are much bigger than previously thought. We estimate that an XXX percent increase in multinational presence via YY percent higher foreign ownership in a given sector, XXX

Finally, a much overlooked aspect in the FDI spillovers literature is the fact that in most firm or plant-level data sets one observes output revenue and total expenditures on variable inputs and hence cannot separate physical output from price. Therefore, it is hard to tell apart a physical productivity improvement from a price effect. This problem is shared by all the productivity literature that estimates productivity at firm and plant level due to lack of firm level prices but it gets worse particularist for FDI and firm productivity literature since markups may respond endogenously to competition effects due to foreign presence, so that output prices charged by domestic firms rise, delivering a positive sign on the estimated spillover coefficient. Our approach is not immune to this concern and the positive and significant coefficient found on our technology-weighted (or not weighted) "knowledge spillover" variable could reflect these type of demand/price changes rather than changes in efficiency due to technology transfer. To address this possibility we follow **?** and compute firm level markups. Interestingly, only forty percent of the change in TFPR induced by increases in foreign investment in technologically close sectors is driven by higher markups. Hence we interpret the rest as improvements in physical productivity due to technology transfer.

Our paper relates to several papers in the literature. The existing "horizontal" spillovers literature is very extensive but fails to find conclusive results regarding the effect of foreign ownership on domestic firms' productivity operating in the same two-digit sector. For example, **?** and **?** find positive horizontal spillovers in the UK and the US, respectively; while **?** finds negative horizontal spillovers in Venezuela and **?** and **?** find no effect in Lithuania and Ireland, respectively. Motivated by the lack of consistent positive horizontal spillovers, papers like **?** suggested that positive spillovers from FDI were more likely to be vertical than horizontal in nature, since 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. In other words, spillovers are most likely to take place through backward linkages, that is, contacts between foreign affiliates and their local suppliers in upstream sectors. We show that our knowledge spillover and technology transfer measures are not driven by vertical customer-supplier relations, although we estimate such "vertical spillovers" in our data with magnitudes that are similar to those found by **?**. Therefore, our results should be interpreted as the learning effect described in the horizontal spillover literature by which domestic firms imitate technology advances by foreign firms or hire high skilled workers that once worked for those foreign firms.

The rest of the paper is as follows. XXX

2 Conceptual Framework

Take a domestic firm operating in a particular four-digit industry s_4. "Knowledge spillovers" refer to the activity of foreign firms that operate within the same two-digit industry as the domestic firm but however, are not direct competitors because they do not produce in the same four-digit industry s_4. As opposed to the measures in the vertical spillover literature, "knowledge spillovers" do not necessarily need to rely on supplier-customer relationships that will be proxied using an input-output matrix.[2]

To fix ideas, here is an actual example from our dataset. We focus on Germany as the host country and take the domestic German company Tepper Aufzuege GMBH (www.tepper-aufzuege.de) as our unit of analysis. Tepper manufactures elevators and it is classified in the four-digit subsector 2822 "Manufacture of lifting and handling equipment" which in turn belongs to the two-digit sector 28 "Manufacture of machinery and equipment, n.e.c.". Classified in the same four-digit industry 2822, we find the foreign company Loedige Aufzugstechnik GMBH (www.loedige-aufzuege.com) owned by the Finnish global leader in the elevator and escalator industry KONE OYJ. Both Tepper and Loedige operate within the same four-digit industry, are direct competitors. On the other hand, operating in the same two-digit sector as Tepper and Loedige (i.e., sector 28 "Manufacture of machinery and equipment, n.e.c.") we find the Japanese subsidiary Kubota Baumaschinen GMBH (www.kubota-baumaschinen.de), which is classified in a different four-digit sector: sector 2892 "Manufacture of machinery for mining, quarrying and construction." The traditional horizontal spillover measure used in the literature assumes that Kubota is a competitor of the domestic company Tepper however, each firm operates in a different product market. Nevertheless, Kubota produces excavators and Tepper can benefit from, for example, hydraulic technology advances achieved by Kubota because the products of both companies (hydraulic excavators and hydraulic elevators) share a very similar key component: hydraulic cylinders and hydraulic motors.[3] The company Kubota will be part of our "knowledge spillover" variable. Notice, in this case, Kubota is not a supplier of Tepper based on information linking four-digit sectors through the I-O matrix and therefore, the "knowledge spillover" do not rely on customer/supplier relationships.[4]

The main take away from this example is that the traditional "horizontal spillover" measure would confound the effect of direct product competitors (such as Loedige) and the effect of non-competitors (such

[2] Vertical linkage measures weight the foreign presence in two-digit industry s_2 by the I-O matrix coefficients linking sectors by input usage across the economy.

[3] 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.

[4] According to the I-O matrix information, sector 2822 is not a supplier to sector 2892. In addition, we could not find evidence of a commercial relationship between Tepper and Kubota.

as Kubota) which in turn can have potential opposing effects. By differentiating between competition and knowledge spillovers we start by setting the appropriate product market space that distinguishes between competitors and non-competitors.

It is possible of course that learning takes place even between direct competitors. To get at this effect and to confirm that our knowledge spillovers variable capture technology transfer, we follow **?** who study industry R&D spillovers identifying technology and product market rivalry effects. They distinguish a firm's position in the technology space using information on the distribution of its patenting across technology fields and construct distinct measures of the distance between firms in the technology dimension. Their refined approach recognizes that a firm is more likely to benefit from the R&D activities of other firms that are *close* in the technology space. Similarly, the standard FDI spillover literature assumes that the productivity of domestic firms will be affected by the overall pool of foreign-owned firms in the same sector. To distinguish competition from knowledge spillovers we argue that knowledge transfers will be more likely to take place in sectors where firms are closer in the technology space while competition effects will be captured by the relative importance of foreign market shares. In sum we argue that technology transfers from foreign to domestic firms should be larger among technologically close firms and sectors in both cases of directly competing domestic and foreign firms and also for firms who are not direct competitors.

3 Data

It is becoming common practice for researchers to work with firm-level data sets to study international linkages among firms (CITE HERE BLOOM ET AL CHINA PAPER). It still remains a problem though to have nationally representative cross-country firm level data sets since no supra-national institution collects them for many countries such as national censuses. When it comes to having nationally representative data from several countries which also connects these firms internationally via foreign investment and trade the problem gets worse.[5] The only source of information for such data is either surveys and/or business and tax registries. Many private vendors compile such data sets using these national sources. Bureau van Dyck, Thompson and Reuters, Dun and Brad Street are such companies. Note that, national representation will still depend on the rules and regulations of the particular country in question since many countries do not collect any financial data from small private firms.

We use ORBIS Global database from Bureau van Dyck. The database covers more than 200 countries and over 10 million firms. The unit of observation is the firm for many countries though in some countries data is reported at the establishment/plant level. Firms can be linked to their domestic and foreign

[5] Each national statistical office conducts its own Census survey and while there have been some efforts on the part of the European Central Bank under their Compnet program to collect such data from central banks and business registeries in the European context, due to confidentiality issues information is not disseminated at the firm level.

parents through a unique ID number structure. The database is similar in its granularity for measuring international firm linkages to WorldBase database that is used extensively in global firm organization literature (CITE ANOTHER BLOOM PAPER). The reason why we prefer ORBIS database is the extent of coverage of domestic firms, which is critical for measuring the knowledge spillovers from foreign to domestic firms. In ORBIS data set we can cover up to 70 percent of the real economy in many countries.

ORBIS collects the data from numerous national providers and harmonizes them into an internationally comparable format. ORBIS database provides a consistent time series representative sample for all countries under our analysis for both *private and public firms.* This is a clear advantage when interested in aggregate economic effects, large firms (those with +250 employees) only account for less than one percent of total manufacturing output or employment in the European countries under analysis.

This database has been previously used in the literature (see for example, **?**, **?**, and Hassan xxx, Gopinath et al. (2016)) however, this paper is first to use foreign ownership information together with information on domestic firms, where domestic firms are a representative sample during the period 1999-2008. The ownership module of ORBIS is particularly burdensome on the time series dimension. Access to ORBIS online or the disk access will only provide ownership information as of the current date. Therefore, to have information that varies over time, it is necessary to download the ownership structure of each firm from each disk.[6]

We focus on a sample of six advanced European countries during the period 1999–2008 for which we have a nationally representative sample of firms with comprehensive information on the variables needed to compute total factor revenue productivity in the manufacturing sector; namely, Belgium, Spain, Finland, France, Italy and Norway.[7] We exclude micro enterprises (those with less than ten employees according to the European Commission definition) since we believe from a theoretical point they are less likely to be subject to technology spillovers from foreign investors. The final sample of firms with ten or more employees accounts for more than ninety percent of manufacturing output in the countries of study and avoids potential concerns about differences in the size distribution of firms across countries. Furthermore, the advantage of focusing on the sample with more than ten employees is that we can directly compare our results to previous studies based on manufacturing surveys since this 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, as well as, the complementary databases we use to construct the

[6]See **?** for a detailed explanation on how to construct nationally representative firm level financial data from BvD products as well as the ownership structure.

[7]Data for Norway corresponds to the period 2000-2008 to work with a nationally representative sample over the full period.

spillover variables that are defined in section ??. More details on the cleaning process and firm level statistics are provided in the Appendix.

3.1 Firm Productivity

Firms' 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 therefore, we measure real output, Y_{it}, as nominal value added divided by a sectoral price deflator. All monetary values are deflated using Eurostat two-digit industry price deflators.[8] Labor input, L_{it}, is measured with the firm's wage bill (deflated by the same two-digit industry price deflator) to proxy for differences in the quality of the workforce across firms. Finally, we measure the capital stock, K_{it}, with the book value of fixed assets and 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 log of labor input, k_{it} is the log of capital input, ω_{it} is the log 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 the output elasticities with respect to inputs in equation ?? is that the firm observes its own productivity ω_{it} at the time of making input choices therefore, rendering the OLS estimates of β_ℓ and β_k inconsistent. We follow the approach suggested in ? that builds on previous work by ? (OP) and ? (LP) and that addresses the concerns raised by ?.[10] ? argue that if the flexible labor input is chosen as a function of unobserved productivity, then the coefficient on the labor input is not identified in the approaches proposed by OP and LP.

We estimate the production function by country and two-digit sector (Table ?? in the Appendix ??

[8] Except for Norway and France that do not have good coverage of industry price deflators at the two-digit and we use the total manufacturing industry price deflator.

[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.

[10] See Appendix ?? for a detail description of the estimation procedure.

shows the estimated elasticities) and winsorize the resulting distribution at the 1 and 99 percentiles by country. Our main results are also qualitatively robust to estimating productivity from a translog production function following the methodology proposed in **?**.

3.2 Firm Foreign Ownership

The ownership section of ORBIS-AMADEUS 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* (fo) 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.

3.3 Firm Patents

We use the number of patents granted to firms to explore the impact of foreign investment on the innovation activity of domestic firms. We use the patent data provided by **?** and match to our sample of firms. **?** identify the firms in AMADEUS that filed at least one EPO (European Patent Office) patent since 1978 and provide the number of granted patents by firm. Patents are dated by application year to measure the formal invention year of the patent (see **?** 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 theirs and assume that the firms that we cannot match were not granted any patents. Notice the only firm level information regarding the patent activity of our sample of firms that we enjoy is the number of granted patents, we do not have any information on the characteristics of each individual patent like the patent activity cluster and we do not have information on the R&D expenditure at the firm level.

[11] Countries have different rules for when the identity of a minority owner needs to be disclosed; for example, 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 its ownership data from the 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, fo for this company is 60 percent.

3.4 Sector Technology Closeness

We use technology closeness measures provided in **?**. They use the U.S. firm level accounting data (sales, employment, capital, R&D expenditure etc.) and market value data from Compustat over the period 1980–2001 and match this information with the 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 basic technology closeness measure which is calculated as the uncentered correlation between all firm i, j pairings as described in Appendix **??**. We use this information on the technology closeness between firm pairs and information on the firms' R&D expenditure to compute bilateral four-digit sector level measures that proxy for technology closeness between four-digit industries (see next section and appendix **??** for a detailed description on how these measures are constructed). The number of four-digit industries covered in **?** is lower than the number of four-digit industries available in our dataset and therefore, we end up with a final sample of 570,000+ firm-year observations.[13] Table **??** in the Appendix shows that there are no major differences in company size across the original sample and the merged sample. Notice the final number of sectors available is dictated by the sectors available in the **?** but it does not mean that we select sectors with positive technology links. Once we aggregate the dataset in **?** to the sector level there can be sectors with no technological links. Table **??** in the Appendix reports summary statistics for the main variables and the final sample of firms used in the analysis. Table **??** shows the correlation.

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-output tables have been 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 information on the US input-output table from the Bureau of Economic Analysis (BEA). VADYM CAN YOU DESCRIBE IN HERE THE STEPS YOU FOLLOW WITH THE US IO TABLE?

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

[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.

Notice we follow the finance literature (see **?** and **?**) and use the US as a benchmark to construct sectoral measures of technology closeness and vertical integration. We assume the six advanced European countries of our sample are close to the US in terms of technology. The US is the leading innovator worldwide and therefore, we can think of the US technology space as the frontier. The advantage of using US data is that we avoid possible endogeneity concerns, technology closeness is likely to be similar in Europe and any deviation could be endogenous to foreign ownership. Similarly, by using time invariant measures of the technology closeness that pre-date the sample of analysis we avoid further endogeneity concerns. And finally, identification does not require that the technology closeness measures of European sectors are exactly the same as those in the US but rather that the ranking of sectors remains stable across countries which we believe is certainly plausible. In any event, by using the US IO structure and US sector technology closeness if anything we are introducing random error in the measurement of our regressors and therefore, reducing the probability of finding statistically significant results.

4 Methodology

4.1 Calculating FDI Spillovers

Horizontal Spillovers. The basic horizontal spillover measure in the FDI literature computes the share of foreign output in total two-digit industry output:

$$\text{HORIZONTAL}_{s2,t} = \frac{\sum_{i\in s2} \text{fo}_{i,t} \times \text{go}_{i,t}}{\sum_{i\in s2} \text{go}_{i,t}}, \tag{3}$$

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

Knowledge Spillovers.

We argue that the two-digit industry classification used in equation **??** is too broad and confounds the effect of competitors and non-competitors. Therefore, we propose to breakdown the HORIZONTAL measure in two:

$$\text{COMP}_{s4,t} = \frac{\sum_{i\in s4} \text{fo}_{i,t} \times \text{go}_{i,t}}{\sum_{i\in s4} \text{go}_{i,t}}, \qquad \text{KNOW}_{s4,t} = \text{HORIZONTAL}_{s2,t} - \frac{\sum_{i\in s4} \text{fo}_{i,t} \times \text{go}_{i,t}}{\sum_{i\in s2} \text{go}_{i,t}}, \tag{4}$$

where $s4$ indicates a four-digit sector classification and the rest of the notation is identical to that in equation **??**. Exploiting information on the four-digit industry classification of firms, these two equations break the aggregate $\text{HORIZONTAL}_{s2,t}$ variable into two components: $\text{COMP}_{s4,t}$ 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 Weighted Knowledge Spillovers.

We enhance our spillover measure with measures of sector technological proximity. To obtain proxies of sectors technology closeness we rely on the work by **?**. 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 comes in as 426 different technology classes. They start by determining the average share of patents per firm in each such technology class over the period 1970 to 1999. The available firm level patent information allows them to define for each firm i the vector of i's technological activity $\mathrm{t}_i = (\mathrm{t}_{i1}, \mathrm{t}_{i2}, ..., \mathrm{t}_{i426})$, where t_{ix} is the share of patents of firm i in technology class x. Then, for each firm pair i, j in their sample, **?** construct measures of technology closeness denoted tech_{ij}. BSvR follow **?**, so that their measure of technology closeness is the uncentered correlation of patent share vectors t_i and t_j:

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

BSvR observe the R&D expenditures by their in-sample firms, labelled r_{jt}, and go on to measure if firm i is impacted by the R&D of other firms $j \neq i$, where the latter is weighted by the measure of technological closeness tech_{ij}. That is, they examine if $\mathrm{spilltech}_i = \Sigma_j \mathrm{tech}_{ij} r_j$ is a determinant of firm i outcomes. BSvR uncover large effects of R&D on related firms.

We follow their notation, except that we use a small-cap font for firm-level variables for contrast with the country-sector variables that we define next.

We hypothesize that knowledge spillovers from foreign firms are particularly strong for firms that operate in technologically close sectors. We aggregate their firm level measures of technology closeness to the sector level and argue that the technology closeness space in the U.S. is similar to that in the set of developed European countries that we are exploring. In particular, we take the following steps to construct sector level measures of technology closeness across sectors. First, for each four-digit sector pair we compute the weighted sum of the technology closeness of firms operating in sector pairs $s4$ and $\tilde{s}4$ as:

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

where $s4$ denotes a particular four-digit industry and $\tilde{s}4$ refers to the rest of four-digit industry sectors in the economy other than $s4$. We use script capital letters for sector-level variable. r refers to the firm level R&D expenditure and firm-level technological proximity tech_{ij} is defined in equation ??.[15]

Next, we argue that the positive knowledge transfer effects from FDI should be larger in sectors that are closer in the technology space. Therefore, we augment 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:[16]:

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

where $s2(s4)$ denote the 2-digit sector that includes the 4-digit sector $s4$. The denominator in equation **??** is chosen such the weight sum to unity in equation **??**. 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 therefore the weights are therefore designed to reflect the relative technological closeness within 2-digit sectors. We define $\text{WTECH}_{s4,\tilde{s4},t}$ only for $\tilde{s4}$ sectors in $s2(s4)$. and then use these weights to compute:

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

Notice that the summation is done over the four-digit industries $\tilde{s}4$, other than a given four-digit sector $s4$ but only within the two-digit sector of $s4$. The WTECH weights are normalized so that they sum to unity for each sector $s4$ in order to retain the interpretation of an increase in the variable $\text{KNOWTECH}_{s4,t}$ as being proportional to an across-the-board increase in the fraction of foreign ownership.

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

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

[15] Notice both the technology closeness measure and the the firm level r expenditure are time invariant. We use the average r expenditure for each firm over time.

[16] Intuitively, the effect of technology transfers from the multinationals in a particular sector, $\tilde{s}4$, which is technologically linked to domestic firms in sector $s4$ might be small, if the sector $\tilde{s}4$ has low output relative to all other sectors which are technologically linked to the sector $s4$.

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

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

Notice in this case, the COMPTECH measure weighs 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.

In what follows we provide an overview of the variables previously described. Figure **??** that shows the relative importance of each type of spillovers according to the OECD classification of manufacturing industries into technological categories.[17] The importance of KNOWTECH (the blue bars) is higher in sectors classified as high-technology. A potential concern could be that the variable KNOWTECH does not contain information different from the variable KNOW (the red bars). Figure **??** in panel (a) shows that while the correlation among the two variables is high, there is new information in the variable KNOWTECH not captured by KNOW. Table **??** in the appendix provides summary statistics for all variables used in the analysis.

5 Empirical Framework

Traditionally, the literature on FDI spillovers has focused on single country studies and estimated an equation of the following type for the sample of domestic firms:[18]

$$\log\left(\text{TFPR}_{\text{i,s2,t}}\right) = \beta\text{HORIZONTAL}_{s2,t-1} + \alpha_i + \delta_t + \epsilon_{i,s2,t}, \tag{11}$$

where $\text{TFPR}_{\text{i,s2,t}}$ refers to revenue total factor productivity of firm i, in sector $s2$, at time t and $\text{HORIZONTAL}_{s2,t}$ 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 at the sector level and changes in firm TFPR.

There are two main concerns with this specification. First, on the methodological side, we would like to interpret β as the change in firm level TFPR derived from changes in foreign activity in a particular

[17] 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 **??** in the appendix shows a more detailed breakdown by the two-digit industry.

[18] In this context, domestic firms are those that were never acquired by foreign-owned investors over the sample period.

sector. However, there is a clear endogeneity concern if the choice of foreign investment is correlated with sector productivity. The sign of this correlation is not straightforward. On the one hand, certain sectors may be expected to have high productivity growth and such sectors are likely to attract foreign investors. On the other hand, foreigners might buy lemons due to informational asymmetries. The advantage of a multi-country study like ours is that we can account for such global trends by the inclusion of sector-year fixed effects and, therefore, estimate the following equation:[19]

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

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 **??** is as we have been emphasizing so far, the coarse level of aggregation of the spillover variable based on the two-digit industry classification. To address this concern we breakdown the HORIZONTAL spillover variable into COMP and KNOW and estimate the following equation:

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

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 of foreign-owned firms in closely related sectors that are not direct competitors. Notice 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 turn back to it in section **??**.

Finally, to test whether the variable KNOW is a good proxy for technological transfers we augment equation **??** and add the technology weighted knowledge spillover variable to the specification. In addition, we notice that the variable COMP can capture both competition and knowledge transfer effects from direct competitors and therefore, we propose to isolate the product market competition effect by including the variable COMPTECH that proxies for technological transfers of direct competitors. The main specification

[19] Notice "global" in here refers to the sample of countries under analysis.

we estimates is the following:

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

where $\mathrm{TFPR_{i,s4,c,t}}$ 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 should be positive if domestic firms benefit from the presence of foreign investors in closely related technological sectors.

SEBNEM: I EDITED UNTIL HERE

6 Results

In order to relate to previous work, we start with the basic specification previously used in the literature. In Table **??**, we report the results from specification **??** 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 **??** by Ordinary Least Squares (OLS), this is inefficient if the variance of the error terms differs across firms. This might be of special concern in the current study that pools firms across different countries. Therefore, all our results are estimated by two-step feasible GLS.[20]

Table ?? about here

We mimic country-level regressions by including country-year fixed effects (corresponding to time fixed effects in a single country study) and firm fixed effects. In column (1), we do not find this variable significant. However, if we include sector × year fixed effects as in specification **??**, controlling for endogenous FDI flows to "hot sectors," we find that foreign presence in the two-digit sector indeed increases the productivity of domestic firms. The estimated coefficients are semi-elasticities so that the coefficient of 0.050 implies that an increase in industry foreign ownership activity of 10 percentage points increases domestic

[20]The first step estimates the equation by OLS and for each firm the square root of the mean squared residuals is calculated. In the second step, the 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 therefore, we winsorize the lower tail of the weights distribution at 5%. Graphical inspection of a partial correlation plot of the regression revealed that there are no obvious outliers. In addition, similar results were found if weights were obtained with a parametric model of the error variance (i.e., estimating standard errors as a function of firm characteristics).

firms' productivity by 0.5 percent. We argue that even controlling for foreign investment selection into particular sectors as we do in column (2), the HORIZONTAL spillover variable constructed at the two-digit industry level confounds two potential countervailing effects derived from competiton and knowledge spillovers. In column (3), we report results from estimating specification **??** using our measures of foreign presence within the four-digit sector and within the two-digit sector but outside the four-digit sector. We find an insignificant effect within the four-digit sector but importantly, the presence of foreign firms in the two digit sector that are not direct competitors results in positive spillovers.

The findings in Table **??** show a positive productivity effect of foreign owned companies operating within the same two-digit industry but that are not direct competitors (i.e., classified in different four-digit industries as the domestic firm considered). The main point is that HORIZONTAL spillover measures constructed at the two-digit industry level combine the effects of direct competitors and firms who are not direct competitors and therefore, it is crucial to start by identifying the appropriate product market space when exploring spillover effects. Next, we provide evidence that the KNOW effect is driven by technological (knowledge) transfers.

In Table **??**, the first column is repeated from Table **??** for convenience, while the other columns explore whether in specification **??** our technology-weighted measures are better able to capture technology spillovers than those previously used. In column (2), we include the KNOWTECH measure and this measure renders the basic non-weighted measure insignificant although the coefficient to the technology-weighted measure is positive significant with the size similar to the value found in column (1). This implies that as expected, the variable KNOW captures technology transfer effects from firms who are not direct competitors. An across-the-board increase in foreign ownership has the size predicted in the previous table; however, as we spell out in more detail below, the impact will be larger when FDI is concentrated in sectors with technological spillovers and lower if not, which could explain why findings 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 the technology space.

In column (3), we include the COMPTECH measure of technology-weighted foreign presence in the same four-digit sector. Within a four-digit sector, foreign ownership may have both competition and knowledge spillover effects even for domestic and foreign firms who are direct competitors, and if that is the case, we expect the unweighted COMP measure to mainly capture the competition effect and the COMPTECH measure to capture knowledge spillovers. And, indeed, the latter measure takes a large positive significant coefficient of 0.10, while the coefficient to COMP turns negative significant, consistent with increased competition hurting domestic firms. The message is that foreign presence at the four-digit level has highly (statistically) significant competition and knowledge spillover effects with knowledge spillovers from firms that are close in technology space being of high economic significance. In addition, as we hypothesized, foreign presence within the same two-digit industry but outside the four-digit has a significant positive impact on the productivity of domestic firms based on knowledge spillovers and no significant competition effects. We have shown that the knowledge spillover variable KNOW has no more information than the technology weighted spillover variable KNOWTECH that captures technology transfer from non direct competitors and therefore, column (4) excludes from the specification the variable KNOW without substantial changes to magnitudes of the estimated coefficients.

Table ?? about here

Table **??** 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. For example, **?** and **?** find positive horizontal spillovers in the UK and the US, respectively; while **?** finds negative horizontal spillovers in Venezuela and **?**, and **?** find no effect in Lithuania, and Ireland, respectively. It is now clear, based on the results described in Table **??**, that the coarse industry aggregation previously used, masks important heterogenous effects.

6.1 Predicted Effect of Increased Foreign Ownership and Aggregate Implications

We consider the effect of an across the board change in fo_i of magnitude Δfo. (Some firms already have 100% foreign ownership, but our prediction should be regarded as the effect of an across-the-board change in foreign ownership, widely distributed across firms.) The predicted effect is

$$\begin{aligned}\Delta \log\left(\text{TFPR}_{\text{i,s4,c}}\right) &= \hat{\beta}_1 \Delta\text{COMP}_{s4} + \hat{\beta}_2 \Delta\text{KNOW}_{s4} \\ &\quad + \hat{\beta}_3 \Delta\text{COMPTECH}_{s4} + \hat{\beta}_4 \Delta\text{KNOWTECH}_{s4}\end{aligned} \tag{15}$$

The change in the right hand side measures are:

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

which implies

$$\Delta\text{COMP}_{s4} = \Delta\text{fo}.$$

An across-the-board change of fo will imply a change in KNOW of

$$\begin{aligned}\Delta\text{KNOW}_{s4} &= \Delta\text{HORIZONTAL} - \frac{\sum_{i\in s4}\Delta\text{fo}\times\text{go}_i}{\sum_{i\in s2}\text{go}_i} && (16)\\ &= \frac{\sum_{i\in s2}\Delta\text{fo}\times\text{go}_i}{\sum_{i\in s2}\text{go}_i} - \frac{\sum_{i\in s4}\Delta\text{fo}\times\text{go}_i}{\sum_{i\in s2}\text{go}_i}, && (17)\\ &= \frac{\sum_{i\in s2, i\notin s4}\text{go}_i}{\sum_{i\in s2}\text{go}_i}\Delta\text{fo}\,. && (18)\end{aligned}$$

I assume here, as everywhere in this draft, that we will replace the time-varying $\text{go}_{i,t}$ weights with average values of go_i The intuition here is that the KNOW effect will be larger for a given 4-digit sector $s4$, if there is a large fraction of the output in the 2-digit sector which is outside the 4-digit sector and, conversely, if some 4-digit sector makes up most of its 2-digit sector, there is less scope for knowledge spillovers. This is why there is a factor of proportionality (but if follows mechanically by putting the in Δfo.

The change in COMPTECH will be

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

and finally, the change in KNOWTECH will be

$$\Delta\text{KNOWTECH}_{s4} = \sum_{\substack{\tilde{s4}\neq s4\\ \tilde{s4},s4\in s2}}\text{WTECH}_{s4,\tilde{s4}}\times\Delta\text{COMP}_{\tilde{s4},t} = \sum_{\substack{\tilde{s4}\neq s4\\ \tilde{s4},s4\in s2}}\text{WTECH}_{s4,\tilde{s4}}\times\Delta\text{fo}\,. \qquad (19)$$

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

$$\Delta\log(\text{TFPR}_{\text{s4}}) = \left(\hat{\beta}_1 + \hat{\beta}_2 * \frac{\sum_{i\in s2, i\notin s4}\text{go}_i}{\sum_{i\in s2}\text{go}_i} + \hat{\beta}_3 * \text{WTECH}_{s4,s4} + \hat{\beta}_4 * \sum_{\substack{\tilde{s4}\neq s4\\ \tilde{s4},s4\in s2}}\text{WTECH}_{s4,\tilde{s4}}\right) * \Delta\text{fo}\,. \quad (20)$$

And then we take the average over the four digit sectors.

And then, I think, the coolest will be second prediction where we set $\Delta\text{fo}_i = 2 * \Delta\text{fo}$ in the half of the sectors that have the large technology spillovers (I need to think a little bit on how to write this, so not tonight) and $\Delta\text{fo}_i = 0$, in the sector that have the lowest spillovers so we can make an interesting point about it is not just how much, but also where, FDI takes place that matters for spillovers.

6.2 Direct Evidence on Technology Spillovers: the Effect on Patenting

Comment on the patent results

Table ?? about here

6.3 Technology or Demand?

The dependent variable is revenue total factor productivity and therefore, we have an efficiency measure that combines the influence of technology and demand factors. Without any further analysis it is impossible to know whether the positive KNOWTECH and COMPTECH effects found in column (4) of Table **??** reflect changes in domestic firms' technology or demand improvements. To see this, lets express revenue total factor productivity as:

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

where P_{it} refers to the firm output price and A_{it} is physical productivity. Ideally, using firm specific prices to deflate nominal output will result in the desired efficiency measure of physical productivity (A_{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. Consequently, changes in TFPR can be the result of changes in physical productivity or changes in output prices. Notice, the firm price is in turn the product of the markup μ_{it} and some function of the marginal cost MC_{it}, as it is clear from the second expression in equation **??**. 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:

$$\Delta\text{TFPR}_{it} - \Delta\mu_{it} = \Delta A_{it} + \Delta MC_{it}, \tag{22}$$

where differences between changes in TFPR and the markup can be originated from changes in physical productivity and/or changes in the marginal cost. We follow **?** to obtain firm level markups. **?** noted that under imperfect competition, input growth is associated with disproportional output growth (as measured by the relevant markup). Based on this insight, **?** 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:

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

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 $P_{it}\mathrm{y}_{it}$ is the nominal value added.

As in **?** we choose labor as the flexible input $\mathcal{J}$. The 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.[21] We find empirical estimates of the median markup close to 1.5.

Column (1) in Table **??** shows our main specification with the dependent variable log TFPR.[22] A ten percentage point increase in KNOWTECH increases domestic firms' productivity by 0.2 percent. Our interest now is focused in determining wether the 0.2 percent increase in revenue productivity is due to changes in output prices or physical productivity. Results in column (2) use the log of firm markup as the dependent variable and show that a ten percentage point increase in KNOWTECH only leads to a 0.08 percent increase in domestic firms' markup. Based on the expression shown in equation **??** it is clear that the increase in firm TFPR induced by changes in foreign ownership activity in closely technology related sectors cannot be fully accounted for by increases in markups. The 12 percentage point difference between increases in revenue productivity and increases in markups has to be the result of changes in physical productivity (ΔA_{it}) and/or changes in the marginal cost (ΔMC_{it}).

? express the 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 case differences between revenue productivity and markups will be explained by increases in physical productivity. However, in our case, we cannot rule out the possibility that foreign entry leads to higher competition in the input market and input prices paid by domestic firms do change with changes in sectoral foreign activity. In fact, in column (3) we show that wages of domestic firms increased in sectors with increased foreign ownership activity that are technologically close but not direct competitors.[23] What is more important for the interpretation of the coefficient of 0.020 in column (1) as technology transfer is that increases in input prices correspond to higher marginal costs which in turn translate in lower efficiency and lower TFPR and therefore could not explain the positive coefficient found. Therefore, if anything, higher input prices mean that our estimate is a lower bound of the true effect of knowledge spillovers in technologically close sectors.[24] Finally, to provide additional evidence that corroborates our claim that the spillovers from foreign firms operating in technologically close sectors are technology driven, in column (4) we use as dependent variable the log of the number of patents at the firm level. Increases in foreign investment in close technology sectors is associated with higher patent activity.

Notice the contrast with the results regarding the competition variable COMP. Almost 75 percent (0.024/0.032) of the negative effect of foreign investment on domestic firm productivity is driven by

[21] Notice we estimated the production function at the two-digit industry level and β_L is constant across two-digit industries.

[22] Notice from this point forward we drop the variable KNOW that we have shown is no longer significant after the inclusion of the technology-weighted measure KNOWTECH.

[23] These results are in line with **?** who find that the presence of FDI in the same industry and region has an indirect effect on wages of skilled workers in private firms in China.

[24] It is also possible that there is a price change in the price of materials. For example, **?** studies the linkage effects of MNCs and suggests that MNCs that heavily source from the host country can foster a larger variety of intermediate goods.

domestic firms charging lower markups after increased competition in the same four-digit sector. The 25 percent left could be driven by higher material prices since there are no effects on wages. On the contrary, again in sectors characterized by high technology closeness among firms, domestic firms increase markups after foreign investment increases (COMPTECH), representing close too 70 percent (0.066/0.096) of the change in revenue productivity. However, without changes in wages or number of patents we cannot be sure of the exact mechanism for the remaining 30 percent increase, which could be driven either by lower material prices or improvements in physical productivity that in this case are not necessarily related to hard technology (the effect on patents is insignificant) but could nevertheless be related to soft technology.

To sum up, increases in revenue productivity induced by higher foreign ownership in closely related technological sectors are partly driven by increases in markups however, especially in the case of knowledge spillovers most of the effect on revenue productivity is the result of improvements in physical productivity.

6.4 Vertical Spillovers

We argue that the positive effect on firm revenue productivity identified through the variable KNOWTECH is not necessarily driven by customer-supplier relationships. In order to control for the possibility that results in column (1) of Table **??** are purely driven by an increase in the business activity among foreign firms and domestic suppliers, we construct vertical backward measures at the four-digit industry level. To do so, we rely on the U.S. I-O matrix coefficients for the year 2007 that are available at the four-digit industry level.[25] We follow **?** 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.[26] The lack of positive horizontal spillover effects at the two-digit industry level shifted the focus of the FDI spillovers literature to the role of vertical spillovers, or productivity enhancing effects along the supply chain. These studies are based on the I-O matrix coefficients which for most countries are only available at the two-digit industry level, therefore, capturing inter-**two**-digit industry spillover effects. In addition, we complement the backward spillover measure constructed at the four-digit industry level, by differentiating among 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. In particular:

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

[26] 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 the rest of the literature, and therefore, we drop it from the analysis (see Table **??**).

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

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

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. IO matrix.

Notice the conceptual difference between the traditional vertical spillover measures and our proposed measure of "knowledge spillovers". The vertical spillover measures from the literature capture the relationship of domestic firm i, operating in sector $s2$ with foreign-owned companies operating in the rest of two-digit economic sectors. Our "knowledge spillovers" measure captures the relationship between domestic firm i operating in the four-digit industry $s4$ and foreign-owned companies operating in related sectors (within the same two-digit industry) but not direct competitors (outside the four-digit sector).

Table **??** proceeds by including sequentially the measures of backward spillovers estimated at the four-digit industry level. Column (1) shows our baseline results. Column (2) shows that including a backward measure that accounts for the extent of foreign presence in upstream sectors (within the two-digit sector - excluding the four-digit) does not change the size of the coefficient on the variable KNOWTECH and turns out to be statistically insignificant. This confirms our claim that the knowledge spillovers are not driven by customer/supplier relationships. Column (3) includes a variable that proxies for the extent of foreign owners in other two-digit upstream manufacturing sectors. This measure is closer to the backward measure suggested in **?** and similar to previous studies in the literature, turning to be positive and statistically significant. Most importantly, column (4) shows that our measures of "knowledge spillovers" are not affected by the inclusion of these backward measure variables. While our focus is to explore horizontal spillover effects and according to column (4) it is now clear that results are not driven by customer/supplier relationships, for completeness 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 IO 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 demand and industry concentration effects,[27] compromising the robustness of the backward measure based on supplier relationships although confirming the knowledge technology

[27] We follow **?** 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 IO 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 $s4$.

closeness importance. Overall we can conclude that the knowledge measures are robust.

Table ?? about here

7 Robustness and Extensions

7.1 FDI Spillovers and Firm Entry and Exit

The results shown so far use an unbalanced panel of firms over the period 1999-2008 so that, firms can enter and exit the sample at different points in time. The main obstacle we face to study firm dynamics is the correct identification of new entrants and exitors. We need to differentiate "true" entry and exit from entry and exit into the sample (i.e., firms that have been operating for a longer period but are only included in the panel at a later point in time because coverage improved or information became available; or firms that no longer form part of the sample because they stop reporting.). 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 exitors 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 this definitions results in columns (1) and (2) of Table **??** are very informative. Column (1) shows that foreign activity is not correlated with the probability of new firms entering the markets while competition from foreign firms in the same product market space increases the likelihood of domestic firms exiting. In fact, results in column (3) focus on the permanent sample of firms and shows that our results are not driven by entry and exit of firms.[28]

Table ?? about here

7.2 Long-Term Effects

A separate concern highlighted in the literature is the possibility that spillover effects take time to materialize and different results should be expected between levels and growth rates. Columns (1), (2) and (3) in Table **??** repeat our main specification estimated in two-, three- and four-year differences.[29] While spillovers within the same four-digit industry are present at short and longer term horizons, becoming if anything more relevant over time, knowledge spillovers from non-competitors take longer to be detectable in the data. Only after three years we start observing positive knowledge spillovers from foreign firms operating in technologically close sectors that do not compete in the product market.

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

[29] Notice the difference specifications do not include firm fixed effects.

Finally, we explore the growth effects that foreign firms have on domestic sales. Columns (4), (5) and (6) repeat the estimation using *log* of output as the left hand side. Results, corroborate our TFPR findings.

Table ?? about here

8 Conclusion

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

Dependent Variable: log Firm Revenue TFP
Sample: Domestic Firms

	(1)	(2)	(3)
$\text{HORIZONTAL}_{s2,t-1}$	0.7	5.0***	
	(0.9)	(0.8)	
$\text{KNOW}_{s4,t-1}$			2.7***
			(0.8)
$\text{COMP}_{s4,t-1}$			-0.5
			(0.3)
Observations	342,832	342,832	342,832
Firm FE	✓	✓	✓
Country-Year FE	✓	✓	✓
Sec2-Year FE		✓	✓
Cluster	cs2y	cs2y	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 $\text{TFPR}_{\text{i,t}}$). $\text{HORIZONTAL}_{s2,t-1}$ stands for horizontal spillovers and it is proxied by share of foreign output in total two-digit sectoral output. $\text{KNOW}_{s4,t-1}$ 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 (see section **??** 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)
$\text{KNOW}_{s4,t-1}$	2.7***	-0.3	1.1	
	(0.8)	(0.9)	(1.1)	
$\text{COMP}_{s4,t-1}$	-0.5	-0.6*	-3.3***	-3.2***
	(0.3)	(0.3)	(0.5)	(0.5)
$\text{KNOWTECH}_{s4,t-1}$		2.6***	1.6**	2.0***
		(0.5)	(0.5)	(0.4)
$\text{COMPTECH}_{s4,t-1}$			10.0***	9.6***
			(1.2)	(1.2)
Observations	342,832	342,832	323,730	323,730
Firm FE	✓	✓	✓	✓
Country-Year FE	✓	✓	✓	✓
Sec2-Year FE	✓	✓	✓	✓
Cluster	cs4y	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 $\text{TFPR}_{\text{i,t}}$). $\text{KNOW}_{s4,t-1}$ 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 **??** 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)
$\text{KNOW}_{s4,t-1}$	1.3**	0.4	0.3	
	(0.5)	(0.6)	(0.7)	
$\text{COMP}_{s4,t-1}$	-0.3	-0.3	-0.4	-0.4
	(0.2)	(0.2)	(0.3)	(0.3)
$\text{KNOWTECH}_{s4,t-1}$		0.7**	0.9**	1.0**
		(0.3)	(0.4)	(0.3)
$\text{COMPTECH}_{s4,t-1}$			0.4	0.3
			(0.7)	(0.7)
Observations	342,832	342,832	323,730	323,730
Firm FE	✓	✓	✓	✓
Country-Year FE	✓	✓	✓	✓
Sec2-Year FE	✓	✓	✓	✓
Cluster	cs4y	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 of the number of granted patents to firm i at time t ($\log(Patents + 1)$). $\text{KNOW}_{s4,t-1}$ 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 **??** 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)	log(μ)	log($wage$)
	(1)	(2)	(3)
KNOWTECH$_{s4,t-1}$	2.0***	0.8*	1.7**
	(0.4)	(0.4)	(0.6)
COMP$_{s4,t-1}$	-3.2***	-2.4***	0
	(0.5)	(0.4)	(0.5)
COMPTECH$_{s4,t-1}$	9.6***	6.6***	2.1
	(1.2)	(1.0)	(1.5)
Observations	323,730	323,730	236,409
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. In column (1), the dependent variable is log revenue firm level productivity at time t (log $\mathrm{TFPR}_{i,t}$). In column (2) the dependent variable is log firm markup (log $\mu_{i,t}$). In column (3) the dependent variable is log firm wage, where wage is computed as the ratio of cost of employment over number of employees. 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 **??** 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)	(6)
KNOWTECH$_{s4,t-1}$	2.0***	1.9***	2.0***	1.9***	2.3***	2.1***
	(0.4)	(0.5)	(0.4)	(0.5)	(0.4)	(0.5)
COMP$_{s4,t-1}$	-3.2***	-3.2***	-3.2***	-3.3***	-3.2***	-3.0***
	(0.5)	(0.5)	(0.5)	(0.5)	(0.5)	(0.5)
COMPTECH$_{s4,t-1}$	9.6***	9.9***	9.7***	10.0***	10.3***	12.1***
	(1.2)	(1.2)	(1.2)	(1.2)	(1.2)	(1.2)
BACKWARD_WITHIN2$_{s4,t-1}$		0.4		0.5	0.5	0.4
		(0.7)		(0.7)	(0.7)	(0.8)
BACKWARD_OUT2$_{s4,t-1}$			1.7**	1.7**	1.5**	0.6
			(0.6)	(0.6)	(0.6)	(0.6)
BACKWARDTECH_OUT2					10.6**	12.6***
					(3.4)	(3.4)
log DEMAND$_{s4,t-1}$						0.8***
						(0.2)
log HERFIN$_{s4,t-1}$						-0.5***
						(0.1)
Observations	323,730	322,523	323,730	322,523	322,523	322,523
Firm FE	✓	✓	✓	✓	✓	✓
Country-Year FE	✓	✓	✓	✓	✓	✓
Sec2-Year FE	✓	✓	✓	✓	✓	✓
Cluster	cs4y	cs4y	cs4y	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_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. BACKWARDTECH_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.

Table 6: FDI Spillovers and Selection

SAMPLE: DOMESTIC FIRMS

	(1)	(2)	(3)
Sample	Full	Full	Permanent
Dependent Variable:	Entry	Exit	log TFPR
Estimation Method:	OLS	OLS	GLS
$\text{KNOWTECH}_{s4,t-1}$	0.3	-4.1	3.0**
	(0.5)	(4.5)	(1.0)
$\text{COMP}_{s4,t-1}$	-0.1	6.4**	-5.6***
	(0.5)	(3.2)	(0.8)
$\text{COMPTECH}_{s4,t-1}$	0.0	-11.7	20.7***
	(1.3)	(9.7)	(2.3)
Observations	323,730	323,730	101,875
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. Columns (1) and (2) use the total sample of firms. Column (3) uses the permanent sample of firms (i.e., firms that are observed during the whole sample period). The dependent variable in column (1) is a dummy variable that equals one if the firm was incorporated in year t. The dependent variable in column (2) is a dummy variable that equals one if the firm is 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. The dependent variable in column (3) is the 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 **??** 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 in columns (1) and (2) and obtained by GLS estimation using as weights the square root of the firm mean squared predicted residuals in column (3). *** denotes 1% significance; ** denotes 5% significance; * denotes 10% significance.

Table 7: FDI Spillovers: Long-term effects

Sample: Domestic Firms

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable:	$\Delta^{j=2}\log\text{TFPR}$	$\Delta^{j=3}\log\text{TFPR}$	$\Delta^{j=4}\log\text{TFPR}$	$\Delta^{j=2}\log\text{go}$	$\Delta^{j=3}\log\text{go}$	$\Delta^{j=4}\log\text{go}$
$\Delta^j\text{KNOWTECH}_{s4}$	0.8	1.8**	3.6***	1.5	5.3***	10.8***
	(0.5)	(0.6)	(0.6)	(1.0)	(1.1)	(1.1)
$\Delta^j\text{COMP}_{s4}$	-3.5***	-4.9***	-5.3***	-13.3***	-15.2***	-17.4***
	(0.6)	(0.6)	(0.6)	(1.2)	(1.3)	(1.3)
$\Delta^j\text{COMPTECH}_{s4}$	11.5***	13.4***	16.1***	45.0***	51.1***	59.0***
	(1.7)	(1.6)	(1.6)	(3.2)	(3.3)	(3.1)
Observations	267,104	207,379	157,914	267,104	207,379	157,914
Firm FE						
Country-Year FE	✓	✓	✓	✓	✓	✓
Sec2-Year FE	✓	✓	✓	✓	✓	✓
Cluster	cs4y	cs4y	cs4y	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 in columns (1) to (3) is log revenue firm level productivity at time t (log $\text{TFPR}_{i,t}$). The dependent variable in columns (4) to (6) is log of firm level gross output at time t (log $\text{go}_{i,t}$) $\text{KNOWTECH}_{s4,t-1}$ 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 (see section **??** 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. Columns (1) and (4) are estimated in second differences; columns (2) and (5) are estimated in third differences and, columns (3) and (6) are estimated in third differences. *** denotes 1% significance; ** denotes 5% significance; * denotes 10% significance.

Figure 1: Spillover measures by technology sector in 2007

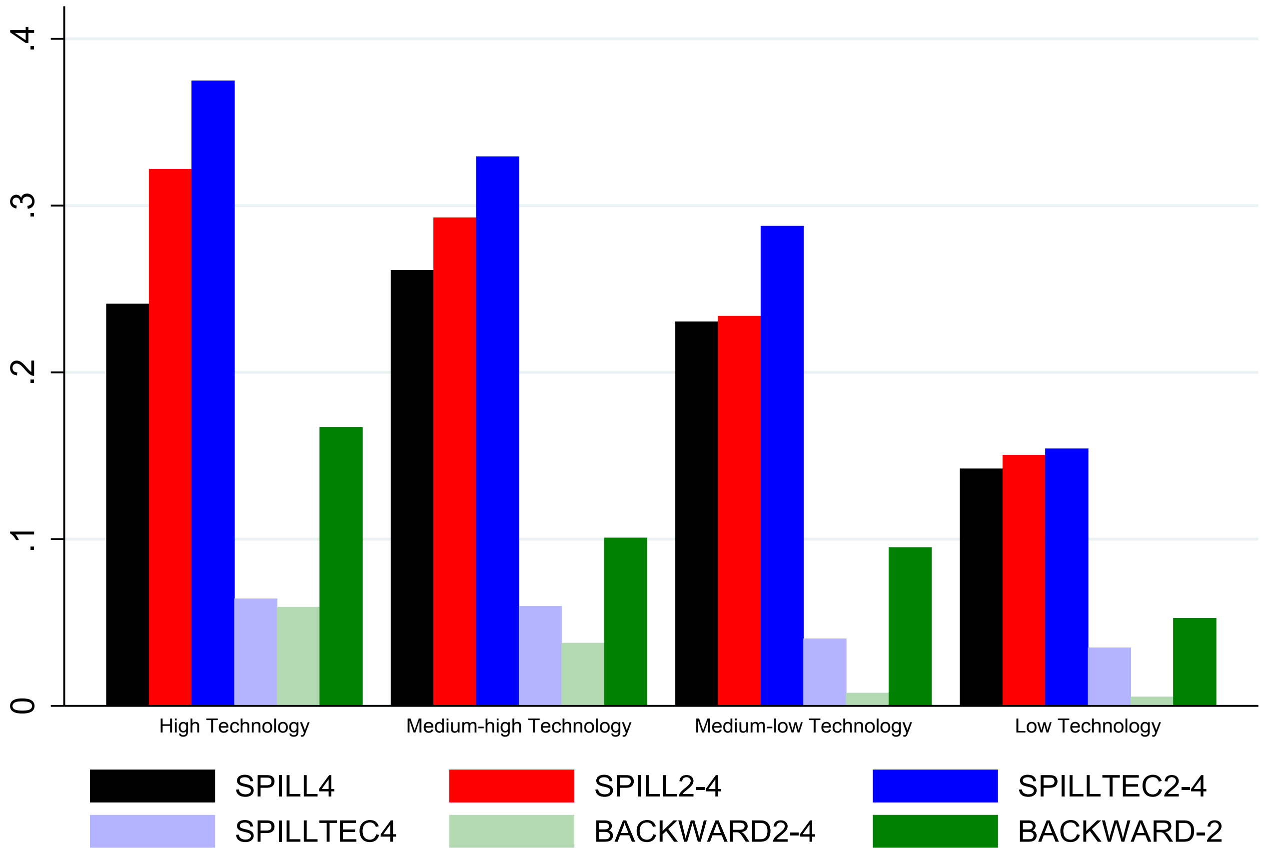
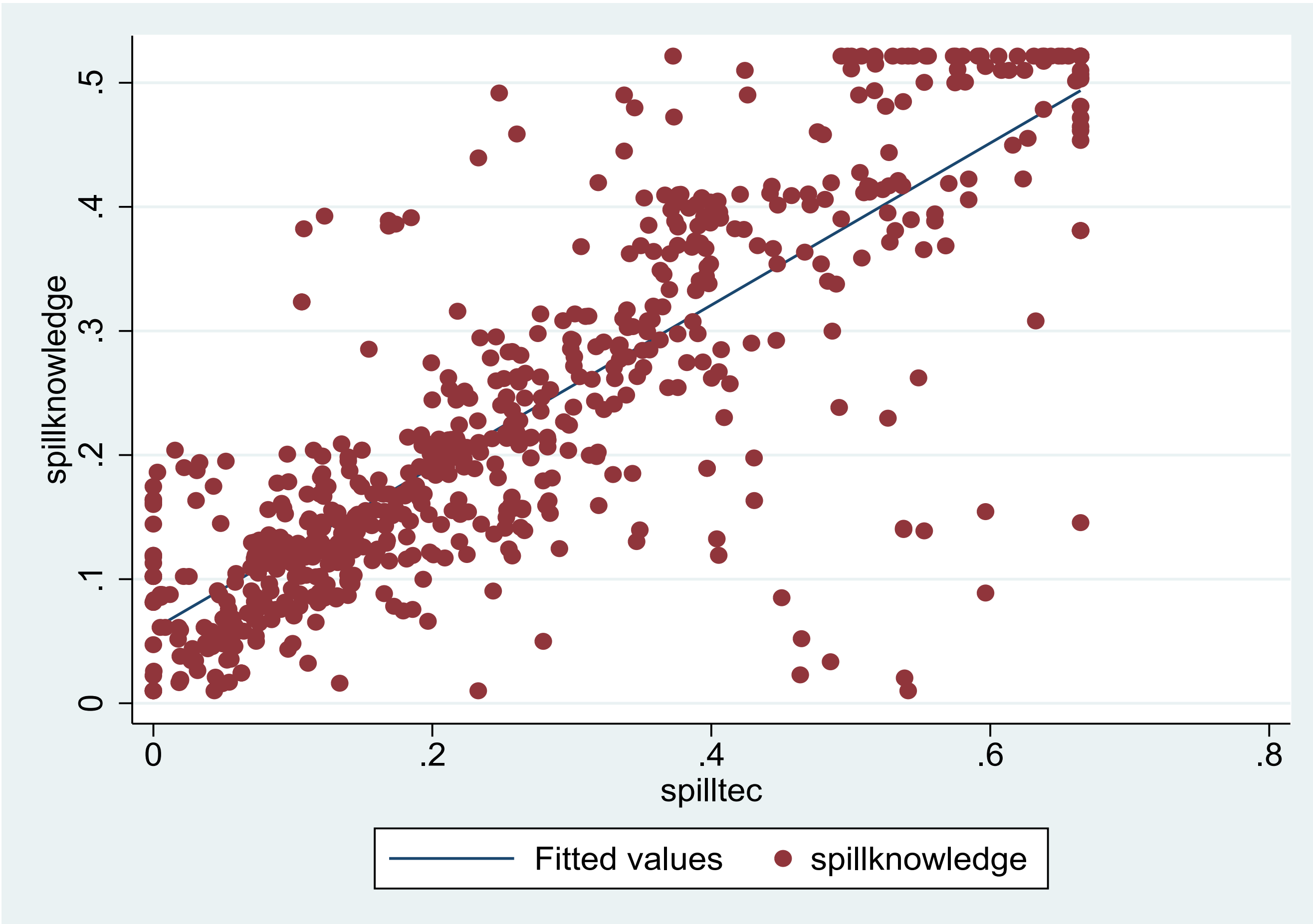


Figure 2: Correlation KNOW and KNOWTECH



Appendix

A Data

- Data sources and sample representativeness both for financials and FO

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 **?** and **?** and further augmented by **?**. **?** (OP) and **?** (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. **?** 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 **?** critique.[30] The following discussion is based on **?**, accommodated to the case of a production functions with two production inputs (see **?** for a general discussion).

For firm i in time period t define:

$$y_{it} = \alpha + \beta_l l_{it} + \beta_k k_{it} + \omega_{it} + e_{it}\,, \qquad \text{A.1}$$

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 $g(.,.)$:

$$\omega_{it} = g(k_{it}, m_{it})\,, \qquad \text{A.2}$$

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

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

[30] **?** 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.

substituting equation (??) into equation (??), we have the following regression function:

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

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 (??) in a form allowing for more lags :

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

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

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

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:

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

Recall from equation(??) 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 (??) gives:

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

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

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

and

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

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 (??) are those outlined in the equation(??), while the orthogonality conditions for equation (**??**) are

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

To proceed with the estimation, we estimate these equations parametrically. In that, we follow **?** and use a third-degree polynomial approximation using first order lags of variable input as instruments.[31]

B.2 Estimation Results

Table **??** reports summary statistics of the output elasticities estimated using the **?** 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 A.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

[31] We use the Stata routine suggested in **?**.

C Appendix Tables and Figures

Table A.2: 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 A.3: 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
BACKWARDTECH_OUT2	519,516	0.29	0.07

Table A.4: 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
BACKWARDTECH_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 A.5: 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
BACKWARDTECH_OUT2	363,617	0.25	0.05	0.16	0.44	0.85	4.18

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). The log of the spillovers variables is computed as $log(variable + 1)$.

Table A.6: Correlation

	log TFPR	HORIZONTAL	COMP	KNOW	KNOWTECH	COMPTECH	BACKWARD_WITHIN2	BACKWARD_OUT2	BACKWARDTECH_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	
BACKWARDTECH_OUT2	0.000	-0.242	-0.070	-0.197	-0.254	-0.078	-0.121	0.143	1

Table A.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)
BACKWARDTECH_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 $\text{TFPR}_{\text{i,t}}$). 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$_{s4,t-1}$ refers to backward vertical spillovers. FORWARD$_{s4,t-1}$ refers to forward vertical linkages. BACKWARDTECH_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 A.1: Spillover Measures by Two-Digit sector in 2007