Università degli Studi del Molise

Dipartimento di Scienze Economiche, Gestionali e Sociali (SEGeS)

ECONOMICS & STATISTICS DISCUSSION PAPER

No. 064/12

Exports vs. foreign direct investment: evidence from cross-country industry data

Filomena Pietrovito

Alberto Franco Pozzolo

Luca Salvatici

Exports vs. foreign direct investments: evidence from cross-country industry data*

Filomena Pietrovito
(Università degli Studi del Molise)

Alberto Franco Pozzolo
(Università degli Studi del Molise, Centro Studi Luca D'Agliano and MoFiR)

Luca Salvatici
(Università degli Studi Roma Tre)

Abstract

The recent process of globalization has been characterized by a rapid increase of foreign direct investments (FDIs), outpacing the simultaneous expansion of arms-length trade (exporting). Trade theory traces back different patterns of internationalization to differences in productivity levels between firms. As in Helpman et al. (2004), we argue that differences in productivity are affected by heterogeneity in firm size. However, we explicitly consider the number of large firms in a sector rather than the size dispersion. Moreover, previous literature performs single country analysis, whereas we extend our analysis to several developed as well as developing countries. By using comprehensive cross-section data on bilateral exports and FDIs (proxied by mergers and acquisitions) over the period 1994-2004, we explain differences across 57 manufacturing sectors in the relative incidence of trade and FDIs. Controlling for other factors affecting the patterns of internationalization and performing several sample splits and robustness tests, our results confirm that sectors with a higher number of large firms are associated with stronger incidence of FDIs relative to trade.

JEL classification: D24, F10, F14, F20, F23
Keywords: exports, foreign direct investments, mergers and acquisitions, large firms

* We acknowledge financial support from the "New Issues in Agricultural, Food and Bio-energy Trade (AGFOODTRADE)" (Small and Medium-scale Focused Research Project, Grant Agreement no. 212036) research project funded by the European Commission. The views expressed in this paper are the sole responsibility of the authors and do not necessarily reflect those of the European Commission. We received helpful comments at the CESifo Workshop on Globalisation, Trade, FDI and the Multinational Firm (2011), at the XVIth International Economic Association World Congress 2011, at the seminar organized by the Department of Economics of University Roma Tre, and at the 12th European Trade Study Group Annual Conference 2010. We thank two anonymous reviewers. A special thanks to the Electronic Resources Area of Bocconi University's library for providing the access to the Worldscope Database. E-mails: filomena.pietrovito@unimol.it; pozzolo@unimol.it; lsalvatici@uniroma3.it. Address for correspondence: Filomena Pietrovito, Università degli Studi del Molise, dipartimento SEGES, via de Sanctis, 86100, Campobasso, Italy.

1. Introduction

An interesting feature of the recent process of globalization is the rapid increase of foreign direct investments (FDIs), outpacing the simultaneous expansion of arms-length trade (exporting). At the end of last century, multinational firms accounted for between two-thirds and three-quarters of world exports and more than one third of world exports were between affiliated firms (UNCTAD, 1999). Since then, global FDIs increased even further, with an expected value of FDI inflows in 2011 around US$1.4 trillion in 2011 (UNCTAD, 2010).

The link between trade and FDIs is strong, since these are possible modes of entering foreign markets. As a matter of fact, firms can serve foreign consumers through two channels: (*i*) produce at home for exports and (*ii*) produce in the destination market through FDIs. The trade literature has shown that what is the best mode of foreign entry depends on the characteristics of the products, firms, sectors and countries involved (Barba Navaretti and Venables, 2004, chapter 6). Similar conclusions have been reached by several strands of firm-level international business research, notably that on the liabilities and benefits of foreignness and that on entry model choices (Slangen et al., 2011).

While there is a long tradition of studies on the factors underlying specific patterns of foreign expansion through trade or FDIs, the literature focusing on measures of relative specialization in trade or FDIs, controlling for the common factors affecting both internationalization strategies, is relatively more recent (Brainard, 1993, 1997; Yeaple, 2003; Helpman et al., 2004; Oldenski, 2010). A well accepted result of these works is that FDIs become more favorable relative to exports as both the size of the foreign market and the costs of exporting increase, and less favorable as costs of setting up foreign production grow (Brainard, 1993, 1997; Yeaple, 2003). More recent contributions, starting from the seminal paper by Helpman et al. (2004), enrich this "proximity-concentration trade-off" taking into account also the role of heterogeneities in firm-level productivity.[1]

While generating important insights, the empirical validation of these studies is still unconvincing, since it is based almost uniquely on analyses focusing on specific countries, for which data on export and outward foreign investment at a disaggregated level are more readily available. In this paper we enlarge the empirical analysis of the trade-off between trade and FDIs using a large dataset including 25 domestic countries, 91 foreign countries and 57 manufacturing industries between 1994 and 2004. Following Helpman et al. (2004), our main focus is on the role of firm size. In

[1] A related issue is the traditional distinction between horizontal and vertical FDIs (see, e.g., Carr et al., 2001). However, this issue is out of the scope of our analysis here.

particular, we make the hypothesis that the presence of a high number of large firms favors foreign entry through FDIs. Our results confirm that sectors with a higher number of large firms are associated with stronger incidence of FDIs relative to trade.

Our analysis contributes to the existing literature along three dimensions. First, we explicitly consider the number of large firms in a sector as a determinant of international commerce between trade and FDIs. Second, we use bilateral flows of trade and FDIs at sector level for a large number of countries. Third, we control for several country- and industry-level characteristics, drawn from the specific literature on trade and on FDIs, that are likely to affect not only the single modes of internationalization, but also their relative incidence. The value added in incorporating industry-level insights in macro-level studies should be particularly high for the international business field, that traditionally has examined internationalization modes separately at the firm and country levels, and should hence have a keen interest in the links between the two (Slangen et al., 2011).

To overcome the limitations of data on bilateral FDIs at the sector level, we use information on the value of mergers and acquisitions (M&As) as a proxy for FDIs. While this is a limitation of our analysis, we believe that it should not affect the qualitative results of our analysis, because cross-border M&As are by and large the most widely used mode of operating a foreign firm (Herger et al., 2008). Moreover, in our robustness checks we provide some evidence that our results are confirmed controlling for the potential impact of greenfield investment.

The rest of the paper is organized as follows. Section 2 briefly discusses the theoretical background and the hypothesis to be tested. Section 3 describes the data used in the analysis. Section 4 presents the empirical model used to test the main hypothesis. The main results of the analysis are presented in Section 5, while Section 6 presents the results of a number of robustness checks. Section 7 draws some conclusions.

2. Theory and hypothesis

The link between exports and FDIs has long been studied in the international business field, where we can trace it back to the eclectic paradigm of Dunning (1977). More recently, in the international economics literature, Helpman et al. (2004) have developed an influential theoretical model to study the impact on the choice between trade and FDIs of a selection mechanism based on productivity, such as that of Melitz (2003).[2] In this framework, firm heterogeneity leads to self-selection in the mode of

[2] In the seminal theoretical model by Melitz (2003), monopolistically competitive firms have different level of productivity,

internationalization, with the most productive firms finding it profitable to meet the higher costs associated with FDIs, firms with intermediate level of productivity serving foreign markets with exports, and lower productivity firms selling only in the domestic market. In the model, a higher within-industry heterogeneity in firm sales is associated with a higher incidence of sales by foreign affiliates relative to exports, because with greater dispersion there is a larger share of firms with a sufficiently high level of productivity to find it profitable to invest abroad.

Using data on exports and on foreign sales of US manufacturing firms in 30 countries and 52 industries, Helpman et al. (2004) also find direct firm-level evidence supporting their theoretical prediction (i.e., multinational firms are more productive than non-multinational exporters) as well as indirect industry-level evidence, since higher firm size dispersion, expressing a higher productivity dispersion, is associated with relatively more foreign affiliates' sales relative to exports. In the same vein, Tomiura (2007) finds that foreign outsourcers and exporters tend to be less productive than the firms active in FDIs or in multiple globalization modes, but more productive than domestic firms, and shows that this productivity ordering is robust when firm size, factor intensity, and sector of economic activity are controlled for. Moreover, Oldenski (2010) extends the analysis of Helpman et al. (2004) showing that greater firm-level heterogeneity in firm size significantly increases FDIs relative to exports also in service industries.

The prediction of Helpman et al. (2004) of a negative relationship between firm heterogeneity and the incidence of trade relative to FDIs critically hinges on two crucial assumptions: (i) that fixed costs to export are lower than those to invest abroad, and (ii) that variable costs of producing abroad are not (much) different from those of producing domestically. In fact, if foreign production was less efficient than domestic production for all firms, for example because of a less skilled labor force, even the most productive firms would find it optimal to export their products rather than to produce them locally. On the other hand, if foreign production costs were lower than the domestic ones only for a subset of the firms population, just the least productive firms would find it optimal to pay the FDIs sunk costs and locate abroad, while the most productive firms would prefer to export (Greenaway and Kneller, 2007). Indeed, Head and Ries (2003) demonstrate that when there are factor price and market size differentials, the ordering of the productivity distribution between multinationals and non-multinationals can be the opposite of that obtained from the Helpman et al. (2004) framework.

---

depending on a draw from an exogenous distribution. With fixed costs to export, only the most productive firms reach a sufficient scale to find it profitable to export. The model is therefore capable of explaining the positive link between productivity and export status, with a causality nexus running from the former to the latter.

In light of the different forces at work, we believe that what effects prevail in shaping the relationship between the level of productivity and the relative incidence of trade and FDIs is an empirical issue. An important difference between our empirical framework and that of Helpman et al. (2004) is that their key explanatory variable is within-industry firm heterogeneity, measured by sales dispersion, while ours is the number of large firms in each country and sector. The choice of Helpman et al. (2004) is a direct consequence of the assumptions made in their theoretical model, namely that firm size depends on the level of productivity, that in turn follows a Pareto distribution. In this setting, the share of large (and highly productive) firms is an increasing function of within-industry firm heterogeneity. However, if firm size followed a different distribution across sectors, for example because of technological factors or economies of scale (Bartelsman et al., 2005), the relationship between dispersion and number of highly productive (large) firms could be non-linear (or even non-monotonic), since sectors presenting similar dispersion measures could feature a different number of large firms. For this reason, we prefer to focus on firm size. Moreover, to focus more explicitly on the role of fixed costs and firm size, we single out the effect of productivity, including a measure of sector level total factor productivity (TFP) as an additional exogenous control in our empirical specification.

Our main hypothesis relates therefore the number of large firms in a sector and the relative specialization in trade or FDIs, and can be stated as follows: *a higher number of large firms in a given sector of a given country is associated with a higher incidence of FDIs relative to trade.*

3. Data and sample[3]

3.1 Dependent variable

A first issue in testing our hypothesis is how to measure the relative specialization in exports or FDIs at the country-sector level. Since we measure FDIs through the value of M&As, we cannot compute the ratio of the value of exports to that of the sales of foreign affiliates, as it is common in the literature (Brainard, 1997; Helpman et al., 2004; Oldenski, 2010). We therefore build a measure of the relative importance of exports on FDIs in the spirit of the literature on revealed comparative advantages (Michaely, 1967; Laursen, 1998), given by the difference between the share of exports in a given sector of a given country with respect to total country exports and the same share for M&As, our proxy for FDIs:

[3] Table 1 lists all variables used in our analysis and their sources.

$$Index_{ij}^{h} = \frac{X_{ij}^{h}}{\sum_{h} X_{ij}^{h}} - \frac{M\ \&\ A_{ij}^{h}}{\sum_{h} M\ \&\ A_{ij}^{h}}. \qquad (1)$$

The first term of our index is the share of exports from country *i* to country *j* in sector *h*, with respect to total exports between the two countries; the second term is the share value of M&As from country *i* to country *j* of sector *h*, with respect to total value of M&As between the two countries. By construction, the index ranges between –1 and +1: it is –1 when sector *h* of country *i* is fully specialized in M&As to country *j*; it is +1 when sector *h* of country *i* is fully specialized in exports to country *j*; the index is equal to zero if sector *h* of country *i* shows the same relative degree of specialization in exports and M&As to country *j*. The index can also be interpreted in terms of similarities between two different entry market modes: 0 indicates the maximum level of similarity (as it is the case in the few instances where the share of exports is equal to the share of M&As); –1 and +1 are opposite cases indicating maximum differences, with M&As prevailing on trade and trade prevailing on M&As, respectively.

Working at a disaggregated level implies the presence of many zero trade and/or investment flows. The index we construct is undefined in the following two cases: (*i*) if total exports (across all sectors) between two countries are equal to zero and/or (*ii*) if total M&As between two countries are equal to zero. In both cases, the denominator of at least one of the two building blocks of our index is equal to zero. To avoid any loss of information, we replace these observations with a value of zero. This choice implies that for couples of countries with no trade or M&As flows across all sectors the index is not centered on zero, i.e., the sum of the indices across *h* is not equal to zero.[4]

By construction, the index in equation (1) ranges between –1 and +1. To avoid being forced to use a truncated regression model, we normalize it taking the following transformation:

$$Index\_norm_{ij}^{h} = \ln \frac{\frac{Index_{ij}^{h} + 1}{2}}{1 - \frac{Index_{ij}^{h} + 1}{2}}. \qquad (2)$$

[4] For a couple of countries that have both trade and M&As in at least one sector, the sum of the index across *h* is by construction equal to zero; to maintain the symmetry of our index we could therefore have assumed a uniform (i.e., *1/h*) trade or M&As distribution. However, since our estimates are not sensitive to this choice, we opted for the more intuitive option of substituting undefined ratios with zeroes.

The normalized index ranges by construction between $-\infty$ and $+\infty$ and it is always defined when the original index is defined (including the zero values).

To construct the index of specialization, we need data on both exports and on the value of M&A operations.[5] The main statistical source of data on exports is the database UN Comtrade, managed by the statistical division of the United Nations, that reports data on the bilateral flows in several industrial sectors. In particular, it contains annual international trade statistics, detailed by commodity and partner country, from 1962 to 2009 for many countries. Commodities are classified according to different recognized classifications, such as the standard international trade classification (SITC) and the harmonized commodity description and coding system (HS). We use the international standard industry classification (ISIC), Revision 3, at 4-digit level to be able to concord data on export with other data used in the empirical analysis.

Data on M&As are sourced from the SDC Platinum *global mergers and acquisitions*, a database provided by *Thomson financial securities data* that records all deals involving a change in ownership of at least 5 per cent of total equity and exceeding 1 million US dollar over the period 1985-2009. The Thomson dataset allows to analyze M&As for a large range of countries and years. This source records two related aspects of cross-border acquisitions: the number of acquisitions and their value.[6] For the purpose of our analysis, and consistently with the literature on M&As, we focus on the value of M&As, and therefore we do not consider undisclosed and incomplete deals for which the value of transaction is not available.

The database also contains information on target and acquirer profiles, such as primary industry and location, that are used in our empirical analysis. In particular, we identify cross-border deals in manufacturing standard industry classification (SIC) codes at 4-digit level.[7]

[5] Both trade and M&As are expressed in current US dollars: it is not necessary to deflate them since our index is constructed as a difference between two shares.

[6] The main sources of information of data on M&As are financial newspapers and specialized agencies like Bloomberg and Reuters. It should be kept in mind that until the mid-1980s Thomson focused very much on M&As for the USA only, and it is only for about the last 20 years that (systematic) M&As data gathering took place for other countries (Brakman et al., 2005).

[7] Domestic M&As, i.e., acquisitions with acquirer and target located in the same country, could still provide access to foreign markets if the target firm is active abroad or if the acquirer is controlled by a foreign firm. However, in the former case we do not know what are the foreign markets (possibly) involved, while in the latter case we have no information about foreign controls: as a consequence, we exclude domestic M&As from our sample.

3.2 Key independent variable

The second issue when testing our hypothesis is how to measure the presence of large firms. First, for each sector we divide the world distribution of firms by total sales in ten deciles. Then, for each sector of each domestic country, we count the number of firms in the first decile of the world distribution of firms by total sales. This indicator proxies for the incidence in each country and sector of those firms that are large enough to overcome the higher fixed costs of expanding abroad through FDIs rather than exports (Helpman et al., 2004).

Data on firm's sales are drawn from the Worldscope database that includes financial statement of about 29,000 companies listed in developed and emerging markets, representing approximately 95% of the global market capitalization. Since we focus on large firms, excluding non-listed companies is unlikely to introduce a relevant bias in our measure of each sector's ability to internationalize. Data are classified according to the SIC classification at 4-digit level.

3.3 Control variables

To avoid omitted-variable bias, we add to the main variable of interest three sets of controls drawn both from the literature on relative incidence of different internationalization modes and from the vast literature focusing on trade and on M&As. First, we control for some relevant sector characteristics in the country of origin. Second, we control for a set of characteristics of the bilateral relationship between each couple of countries. Finally, we include some sector characteristics that are specific of each pair of countries.

Country of origin sector-level variables.

First, we control for average sector wages (expressed in US dollars deflated by using the US consumer price index, with base year in 2000), obtained from UNIDO (Indstat4, 2008 version). Second, following Helpman et al. (2004) who show that capital intensity is a useful predictor of the incidence of exports relative to FDIs, we use data from UNIDO to construct a measure of capital intensity defined as the ratio between capital and number of employees at sector level. Third, following again Helpman et al. (2004) who show that technological intensity favors FDIs relative to exports, we include the number of utility patents granted by the US Patent Office that have been produced worldwide in each sector, provided by the national bureau of economic research (NBER).[8] Finally, as discussed above, we

[8] Since the original data on patents are classified according to the US Patent Classification, we combined them with other

include the average industry TFP, calculated under the assumption of constant returns to scale Cobb-Douglas production function:

$$TFP_i^h = \frac{Y_i^h}{(K_i^h)^{\alpha}(L_i^h)^{1-\alpha}}. \quad (3)$$

where (omitting indices): *Y* is the sector value added, *K* is the stock of capital at the sector level and *L* is the number of employees in the sector, assuming a capital share of 1/3 and a labor share of 2/3.

Total factor productivity at the national sector level was calculated from data on investment and labour from UNIDO (Indstat4, 2008 version), and estimating each sector's capital stock with the inventory method (Bernanke and Gurkaynak's, 2002; Isaksson's, 2007). In particular: (*i*) for each country we calculated the sector's share of investment using flow information for the first five years of data available; (*ii*) we used investment shares to divide information on each country's total capital provided by UNIDO's World Productivity Database across sectors; (*iii*) we used the estimates of the country and sector specific initial stock of capital obtained as described above as the starting point to apply the inventory method, i.e., adding each year's value of real term investment and applying a sector specific rate of depreciation to account for obsolescence.

Bilateral country-level variables.

The empirical literature has identified a large set of variables that influence foreign markets entry modes, though the magnitudes and even the signs of the impact on either trade or FDIs are not always consistent (see, for example, Blonigen, 2005; Disdier and Head, 2008; Helpman et al., 2008; Herger et al., 2008; Oldenski, 2010; Slangen and Beugelsdijk, 2010; Wang et al., 2010; Slangen et al., 2011). Distance directly increases transaction costs because of the transportation costs of shipping products, the cost of acquiring information about other economies, and the cost of finding a partner and contracting at a distance. Similarly, common legal system, common language, common religion, common borders and colonial ties are expected to affect bilateral relationships, both through trade and investment.

Our data on bilateral characteristics (distance, number of islands and landlocked countries in a pair, common language, contiguity and colonial ties) are drawn from the dataset provided by the *centre*

information adopting the correspondence scheme between the US Patent Classification and the International Patent Classification and between the latter and the ISIC3 provided by Johnson (2002).

d'etudes prospectives et d'informations internationales (CEPII).[9] The only exception are the data on common legal systems that are from Djankov et al. (2002).

Bilateral sector-level variables.

We consider two bilateral sector-level variables. First, bilateral trade tariffs with an expected negative sign, since firms shift to FDIs according to the well-known "tariff jumping" effect pointed out in the literature (Brainard, 1997; Carr et al. 2001; Markusen and Maskus, 2002; Yeaple, 2003; Helpman et al., 2004). To make data comparable to other data used in the analysis, we aggregate HS 6-digit level data on tariffs from TRAINS to the 4-digit ISIC classification through simple averages.

Second, building on the results of Chaney (2011) – who show that the existing contacts of a firm can be used to find new ones – we include in our specification a "network index" calculated as the number of common partners in trade and in M&As of each couple of countries (Francois, 2010). We expect that a higher number of common partners in exports (or in M&As) between two countries increases trade (or M&As) specialization between those same countries. Data on the number of common partners is built from our information on trade and FDIs.

3.4 Sample and summary statistics

Matching our different sources, we construct an original database that associates bilateral trade and FDIs flows at sector level in a common classification, for a sample of developed as well as developing countries. Ideally, the full set of industries should be included, with the extent of tradability reflected in transport costs (Brainard, 1997). In practice, however, data on transport costs are only available for industries in which trade exists.

As a consequence, industries including finance and utilities were excluded, along with wholesale and retail trade, because of the non-tradable nature of these activities. We also excluded agriculture and primary sectors (i.e., mining and oil and gas extraction) due to the lack of data on productivity. As a result, we focus on manufacturing sectors i.e., sectors with an ISIC code between 1511 and 3720.

Since our measures of M&As and sales are available in the SIC classification, we made a connection between the manufacturing sectors identified by the SIC code and data classified according to the ISIC code, both at 4-digit level, using the concordances produced by Statistics Canada, as in

[9] The CEPII follows the great circle formula and uses latitudes and longitudes of the most important cities (in terms of population) to calculate the average of distances between city pairs. Data on distances are available at: http://www.cepii.fr/anglaisgraph/bdd/distances.htm. We also adopted distances between capitals as an alternative measure and the results remain unchanged.

Brakman et al. (2005).[10] To take into account that at the 4-digit level of disaggregation we have a large number of empty cells, both in exports and in M&As, we aggregate data available at 3 digits of ISIC classification. Matching different datasets yields data on 25 domestic countries and 91 foreign countries, covering 57 manufacturing industries at the 3 digits ISIC level from 1994 to 2004.

As shown by many theoretical and empirical studies (e.g., Caballero and Engel, 1999), investment dynamics are lumpy. This is even more true in the case of FDIs and M&As (see, for instance, Brakman et al., 2005). For these reasons, although our sample covers 11 years, we estimate our empirical model on data averaged over the entire sample, to smooth time-series variability.

Table 2 presents the descriptive statistics for the variables used in the estimations. It shows substantial variation in all our key variables.

The dependent variable (*Index_norm*) has an average value of 0.051 and a standard deviation of 0.188, with values ranging from -0.884 to 7.058. Positive values are associated with couples of countries presenting higher exports share than M&As share in a given sector, while negative values are for country pairs presenting higher M&As shares than exports shares in a given sector.

Considering our explanatory variables, the number of firms in the first decile of the world distribution of firms by total sales is 3 with a high within sample variability (values range from 1 to 55. The TFP levels (in logs) range from 1.614 to 7.785 (average value: 5.073) and the sectors presenting (on average) the highest values are: Refined petroleum products, Tobacco products, Motor vehicles and Automobiles. The number of patents, reflecting the level of technological development, shows an average value of 17 and a high variability since it ranges between 0 and 1,465.

Concerning bilateral characteristics, tariffs show a high variability, with values ranging between 0 and 58.2 per cent and an average level of 11.7 per cent. The average number of common partners in trade is 58, with values ranging between 0 and 117, whereas the average number of common partners in FDIs is much lower and the range narrower (between 0 and 30). This difference highlights that the two "networks" are quite different and the former is much larger than the latter (consistently with the lower fixed costs assumption, again).

In Table 3 we report simple correlations among the variables used in the empirical model. The correlation between the normalized index and the number of firms in the first decile of the world distribution of sales is negative, suggesting that having a larger share of world large firms favours FDIs relative to trade. Further, TFP levels are positively correlated with the relative importance of exports:

[10] The concordances used are available at: http://www.macalester.edu/research/economics.

higher levels of TFP in a given sector determine higher trade compared to M&As flows between two countries. Higher wages in the domestic country are also positively associated with the incidence of exports, while the contrary is true for capital intensity and patents.

Bilateral correlations are suggestive, but they do not control for potentially confounding factors. For this reason, in what follows we perform a more refined econometric analysis.

4. Methodology

To analyze the underlying motives of the composition of international commerce between trade and FDIs, we design two sets of regression models.

The first set is used to test our hypothesis considering the full set of information available (67,911 observations), including all zero values. Using the normalized index and the three sets of controls defined above, we estimate the following model:

$$Index\_norm_{ij}^{h} = \alpha + \beta_1 Number\_large\_firms_{i}^{h} + \beta_2 Z_i^h + \beta_3 T_{ij} + \beta_4 X_{ij}^h + \\ + \beta_5 DU_i + \beta_6 DU_j + \beta_7 DU^h + \varepsilon_{ij}^h. \quad (4)$$

where (omitting indices): *Index_norm* is the measure defined above of the relative incidence of trade relative to FDIs in sector *h* and countries *i* and *j*; *Number_large_firms* is the number of country *i* firms in the first decile of the world firms distribution of total sales in sector *h*; *Z* is the set of sector specific control variables for the exporting country in each sector (i.e., TFP, wage levels, capital and technological intensity); *T* is the set of control variables describing the bilateral relationship between countries (e.g., distance, common language and common religion); *X* is the set of control variables describing the bilateral relationship between countries in a given sector (i.e., tariffs, number of common partners in trade or FDIs); and *DU* are three sets of dummies controlling for the domestic country, the foreign country and the sector-specific fixed effects.

In our specification, we test the impact of firm size on foreign market entry modes. However, a potential reverse causality problem emerges if the entry mode affects firm size. As a matter of fact, M&As can lead to an increase in the productivity of the bidder, and therefore in its production and sales.[11] In the same vein, foreign trade could increase the firm productivity, for example if it allows the exploitation of economies of scale. We address this potential endogeneity problem instrumenting the number of large firms in each country and sector with two variables: the number of large firms in a

[11] We are indebted to an anonymous referee for pointing out this issue.

given sector *h* in all countries except country *i*; and the number of large firms in a given country *i* in all sectors, except for sector *h*.

As a robustness check, in a second set of regressions, we estimate two Heckman correction models, to separately account for the cases in which there is no trade and/or no M&As in any sector between a couple of countries. In this way, we transform the selection bias problem into an omitted variable problem, which can be solved by including an additional variable, the inverse Mills ratio between the regressors. The Heckman two-step approach allows us to distinguish the impact of preferences on the extensive as well as the intensive margins.

In the first Heckman model, the extensive margin is represented by the probability of any form of internationalization. Accordingly, we estimate the impact of the independent variables included in equation (4) on a binary variable which is equal to one if trade and/or M&As exist, and zero otherwise. In the second step (intensive margin), we estimate the same regression in equation (4), but on a reduced sample of 60,298 observations, excluding all cases in which both trade and M&As are zero and including, among the independent variables, the inverse Mills ratio from the first stage.

Since our database includes only a relatively small number of cases with positive flows of both trade and FDIs, it is also interesting to assess the impact of the presence of large firms on the probability of internationalization through both trade and M&As. Accordingly, in the second Heckman model, the first step estimates the impact of the number of large firms on the probability to enter foreign markets using both trade and FDIs, and the second stage focuses on those cases in which both exports and M&As are present (3,755 observations).

In both Heckman models, identification of the first stage is obtained through the exclusion of the measures of contiguity and colonial ties from the second step estimates.

5. Results[12]

5.1 Baseline regression and sample splitsThe first step of our empirical analysis is the estimation of the model described in equation (4), where the dependent variable is the index of relative specialization in trade or FDIs. We estimate this specification on a sample that includes all the 67,911 cases.

Results in column 1 of Table 4 show that sectors with a higher number of large firms have a stronger incidence of FDIs relative to trade. The negative coefficient of the number of firms in the first

[12] All estimates reported in this section include three sets of dummies controlling for the domestic country, the foreign country and the sector-specific fixed effects, as stressed in section 4.

decile of the world distribution by total sales, statistically significant at the 99% level, confirms our main hypothesis that when the distribution of firms in a given sector-country is shifted towards large firms, it is more likely that the prevailing internationalization mode is direct investment rather than trade.[13] It therefore provides further support to the evidence of Helpman et al. (2004) that larger firms are more likely to be able to afford the higher fixed costs required to invest abroad. Reassuringly, in a number of unreported regressions we have verified that this result is confirmed also using the number of firms in the first quintile of the world distribution by total sales as a threshold to define large firms.[14]

To analyse whether the relationship of interest is affected by some structural features of the domestic sectors, we have then split our sample according to (*i*) the average wage level and (*ii*) the average capital intensity.

Columns 2 and 3 report the results obtained splitting the sample between countries and sectors with average wages above and below the median. In the former sectors, the coefficient of the number of large firms is negative and statistically significant, while in the second it is positive. Our general result of a preference for FDIs over trade when the presence of large firms is higher is therefore driven by the sectors paying wages above the median. Apparently, large firms in sectors paying high wages try to find abroad cheaper labor inputs, consistent with a cost-minimization strategy. On the contrary, those in sectors paying low wages may find it optimal to internationalize only through trade: this determines an increase in the positive value of our index leading to the positive coefficient of column 3.[15]

Along a similar vein, the preference of large firms for an FDI-driven internationalization is due to the firms with a capital intensity ratio above the median, i.e., by more productive sectors (columns 4 and 5).[16] These conclusions are reinforced when we look at the sign of the coefficients of the control variables. Sectors with higher average wages show a lower relative specialization on trade, except for

[13] The coefficient of 0.0116 reported in column 1 of Table 4 implies that an increase of 1 per cent in the number of large firms in a given national sector determines a reduction of the value of our normalized index of specialization (*Index_norm*) by 0.0116 per cent. In turn, starting from a value of the index of zero (i.e., symmetry between trade and FDIs) and from the average number of large firms (3), it implies that one additional large firm determines an increase in the value of the non-normalized index of specialization (*Index*) of 0.19 per cent.

[14] Results are available upon request.

[15] Recall that the second term in our index takes a value of zero for country pairs with no bilateral M&As flows. To control for this undesired consequence of our index, in an unreported regression, available upon request, we have estimated equation (4) on the reduced sample that includes only non-zero observations for both trade and FDIs, controlling for potential sample selection bias through the inclusion of an Heckman correction term (see also section 5.3 below). Reassuringly, we find a negative coefficient for the number of large firms also in the sub-sample of low wage sectors.

[16] Also in these cases, results (available on request) are robust when we use as a threshold the number of firms in the first quintile of the world distribution.

sectors with wages and capital intensity above the median. Moreover, firms in capital intensive industries have a stronger incentive to invest abroad.

The fact that sectors with higher average productivity have a higher relative incidence of trade may come as a surprise. However, it should be noted that the impact is quite small: a 10 per cent increase in productivity leads to a trade share only 0.2 percentage points larger than the corresponding M&As share.[17] Moreover, in an unreported regression, substituting the continuous measure of TFP with a set of four dummies for each quartile level we verify the presence of nonlinear effects of productivity on our index of specialization. The positive and statistically significant coefficient of the dummy for sectors in the top quartiles of the within-country distribution shows that only very high levels of productivity influence the choice between trade and FDIs. In other words, only the most productive sectors tend to favour exports with respect to foreign investment, while in all the other groups the opposite is true. Such a finding is confirmed by the fact that the number of patents produced in a sector show a positive and significant impact on the preference on trade over foreign investment. As a matter of fact, sectors with a high technological intensity, that are likely to be the most productive, prefer to produce at home and then export, instead of producing in foreign countries. This result is only apparently in contrast with the hypothesis that sectors with a higher presence of more productive and therefore larger firms are relatively more likely to invest abroad than to export. Indeed, it is possible that in some sectors, technological constraints make it difficult even for highly productive firms to reach the size that allows to overcome the fixed costs of FDIs, forcing these firms to internationalize through trade.

Regarding country-level bilateral characteristics, most of the control variables related to trade and investment costs (e.g., common language and religion) present a negative and statistically significant coefficient providing evidence that these factors favour FDIs with respect to trade. The opposite is true as far as the distance variable is concerned, and this flies in the face of traditional gravity models predicting that trade costs increasing with distance should promote investment. More recent papers focusing on FDIs, though, provide a set of explanations pointing in a different direction. In Head and Ries (2008), for instance, monitoring requires costs that are increasing in distance between head office and subsidiary. While Slagen et al. (2011) argue that arm's length affiliate sales are likely to decline

[17] Starting from a value of the index of zero (i.e., symmetry between trade and FDIs) and from the average level of *TFP*. In fact, the coefficient of 0.0595 for *TFP* (*log*) in column 1 of Table 4 implies that a 10 per cent increase in the average level of sector TFP determines an increase of our dependent variable (*Index_norm*) of 0.6 per cent and an increase in the value of the non-normalized index of specialization (*Index*) of 0.2 per cent.

with cultural distance, but this is not necessarily the case with arm's length exports, which may in fact increase with cultural distance.

Concerning sector-level bilateral characteristics, the coefficient of applied tariffs is negative and statistically significant, providing evidence of the "tariff jumping" effect: higher tariffs provide an incentive to switch from trade to investment abroad. The coefficients associated with the number of common partners in trade or in FDIs confirm the relevance of the network effects. Apparently, firms in sectors with a higher number of foreign contacts are more likely to enter an additional market, and sectors benefit from the contacts of their contacts. In other words, if a firm *k* has a contact in country *j'* which itself has a contact in country *j*, then firm *k* is more likely to enter country *j*. Furthermore, our results show that the trade and investment contacts form different networks and have opposite impacts on the internationalization choices.

In Table 5 we present the findings obtained considering different samples of countries. First, we consider the choice between different entry market modes made by firms operating in developed countries, distinguishing G-10 (Belgium, Canada, France, Germany, Italy, Japan, Sweden, Switzerland, United Kingdom and United States) and OECD. In this respect, we consider G-10 and OECD as origin countries and, then, consider the group of OECD as both origin and destination countries. Next, we test our main hypothesis limiting the sample to the developing countries as destination markets.

Columns 1 and 2 refer to the internationalization strategies of firms based in G-10 and OECD countries, respectively. The preference of large firms for FDIs is higher with respect to the baseline specification, while the sign and the significance of the other coefficients remains by and large unchanged with a few exceptions. The first exception is represented by the wage coefficient, since developed countries' firms trading is not negatively affected by labor costs (actually, higher wages have a positive impact on trade choices). The second one is due to the fact that trade between most developed countries and their partners is not impaired by common language and religion.

In column 3 we analyze the determinants of foreign market entry modes for the subsample of OECD countries towards other OECD members. Restricting the sample of origin and destination countries does not change the overall picture, but trade costs, both in terms of distance and tariffs, are not significant in explaining the internationalization choices among developed countries.

Considering the group of developing countries as destinations of foreign investment, the overall results are confirmed. Indeed, the coefficient of our variable of interest, that is the number of large firms, is still negative and significant. In other terms, for internationalization toward developing

countries, large firms still prefer M&As relative to trade though with a lower intensity with respect to the baseline estimation.

Finally, we control for the possibility that M&As flows are influenced by the existence of other types of FDIs. In particular, we assess the sensitivity of the results to the presence of greenfield investments, splitting the sample of destination countries according to their ability to attract this type of investments. In practice, we consider the share of the number of greenfield in each destination country over the world greenfield investment and separate the countries with a share of world greenfield investment above the median from those below.[18]

The last two columns of Table 5 report the results for the two groups of countries depending on whether they show a share of world greenfield investment above or below the median. Our baseline result holds in both cases. More interestingly, results for the group of countries attracting a high share of greenfield investment are similar to those for developing countries. This confirms the observation that developing countries tend to attract a relatively large number of greenfield investments (UNCTAD, 2010).[19]

5.2 Robustness checks: Heckman models

In Table 6 we provide some robustness tests of our previous results, by estimating the two Heckman models describer in section 4.

The first step of the first Heckman model consists in estimating the probit model where the dependent variable is the probability of internationalization through trade and/or FDIs.[20] Results, reported in column 1 of Table 6, show that the number of large firms in a given sector exerts a positive impact on the extensive margin, i.e. on the probability of accessing foreign markets through trade and/or FDIs. This confirms the key prediction of selection models *à la* Melitz (2003), namely the existence of a productivity ordering of firms according to their participation in international markets.

In the second step of the Heckman model, we estimate an instrumental variable regression to account for the incidence of our key independent variable on the relative importance on trade on FDIs, by excluding all cases in which both trade and FDIs flows are zero. We include among the regressors the Heckman correction term, calculated from the previous probit regression, to account for the restriction of the sample.

[18] Data on the incidence of greenfield investment is obtained from UNCTAD (http://www.unctad.org).
[19] We are indebted to an anonymous referee for pointing out this issue.
[20] The estimates are in this case obtained from a smaller sample of 63,176 observations since observations are dropped when country or sector dummies perfectly predict the absence of FDIs.

Column 2 of Table 6 shows the results. In terms of our main hypothesis, it is confirmed that a higher number of large firms is associated with a higher incidence of FDIs. This implies that our main results are confirmed also excluding observations with both trade and FDIs flows equal to zero. Comparing coefficients of our key explanatory variable in column 1 of Table 4 and in column 2 of Table 6, it can be inferred that, among internationalized firms, the preference of large firms for FDIs is lower (-0.0099) than in the overall sample (-0.0116). This is consistent with the fact that in the second stage we drop the least productive firms. The positive and statistically significant coefficient of the inverse Mills ratio confirms that the characteristics of sectors featuring positive trade or FDIs bilateral flows are such that they show a higher incidence of trade.[21]

Results on the extensive margin for the second Heckman model, reported in column 3 of Table 6, show that the probability of accessing foreign markets with both exports and FDIs increases when the number of large firms is higher. In the second stage, we analyze the effect of our key explanatory variables on relative trade specialization focusing on those 3,755 cases with positive flows of both exports and FDIs (column 4 of Table 6). In the restricted sample of the most internationalized sectors, the incidence of large firms still augments the preference for FDIs over trade. Interestingly, the exclusion of country pairs with no trade or no M&As does not cause a sample selection bias, as it is confirmed by the statistically insignificant coefficient of the inverse Mills ratio.

6. Conclusions

The firm choice between exporting at arms' length and foreign direct investment, traditionally modeled as a proximity-concentration trade-off (Brainard, 1993, 1997), has been enriched in more recent contributions (Yeaple, 2003; Helpman et al., 2004; Oldenski, 2010) taking into account heterogeneity in firm productivity. While generating important insights, these studies have generally focused on single-country analysis.

In this paper, we study the determinants of the composition of international commerce between exports and FDIs across sectors and countries, explicitly considering the number of large firms in a sector, instead of the heterogeneity in firm productivity. We test the hypothesis that a higher number of large firms leads to a specialization on foreign investment using a novel dataset including 25 domestic countries, 91 foreign countries and 57 manufacturing industries covering the period 1994-2004. We

[21] Since the inverse Mills ratio is an estimated regressor obtained from the probit estimates, we have used bootstrapped standard errors with 100 replications, as in Helpman et al. (2008).

found sound and convincing evidence in favor of this hypothesis, consistent with the predictions of Helpman et al. (2004).

In addition, we are able to shed some light on the empirical linkages between internationalization choices and a rich set of economic variables. First, we confirm the relevance of the well-known "tariff-jumping" rationale for FDIs. Second, while most of the literature studies the role of the host country wages in attracting/repelling FDIs, we focus on the role of the domestic country wages showing that they encourage outbound FDIs, but this is true only for sectors paying lower wages and characterized by low capital intensity. As far as capital intensity is concerned, our results confirm the findings of Helpman et al. (2004) for the US: more capital-intensive sectors export less relative to FDIs. Finally, we provide empirical support to the predictions of the most recent network models (Chaney, 2011) providing a theory of the distribution of entry into foreign markets without any assumptions on firms' productivity distribution. Our results show that export and investment contacts are substitutes rather than complements: being part of a trade network increases the likelihood of using the same mode of internationalization when entering into another foreign market.

Investigating in more detail the characteristics of sectors that are likely to drive our results, we discover that the preference of sectors characterized by a high presence of large firms for an FDI-driven internationalization is due to a specific group of sectors, namely those with an average level of wages and a capital intensity ratio above the world median. Moreover, our results are robust to different country groups splits as well as to the exclusion of different sets of zero trade and/or FDIs flows. In this perspective, the treatment of the zeroes through the Heckman selection model allows to distinguish the impact both on the extensive as well as on the intensive margin. This suggests to extend the present analysis to explain the distribution of the number and the geographic location of foreign markets, putting the emphasis on the extensive margin of trade and FDIs.

References

Barba Navaretti, G., Venables. (2004). "Multinational Firms in the world economy". Princeton University Press.

Bartelsman E., S., Scarpetta, F. Schivardi. (2005). "Comparative analysis of firm demographics and survival: evidence from micro-level sources in OECD countries". Industrial and Corporate Change, Volume 14, 365–391.

Bernanke, B. S., Gurkaynak, R. S. (2002). "Is growth exogenous? Taking Mankiw, Romer and Weil seriously". in NBER Macroeconomics Annual 2001, Volume 16, MIT Press.

Blonigen, B. (2005). "A review of the empirical literature on FDI determinants". *Atlantic Economic Journal* 33: 383-403.

Brakman, S., Garretsen, H., Van Marrewjk, C. (2005). "Cross-border mergers and acquisitions: on revealed comparative advantage and merger waves". CESifo Working Paper n. 1602.

Brainard, S.L. (1993). "A simple theory of multinational corporations and trade with a trade-off between proximity and concentration". NBER Working Paper n. 4269.

Brainard, S.L. (1997). "An empirical assessment of the proximity-concentration trade-off between multinational sales and trade". *American Economic Review* 87:520-44.

Caballero, R.J., Engel, E.M.R.A. (1999). "Explaining investment dynamics in US manufacturing: a generalized (S, s) approach". *Econometrica* 67:783-826.

Carr, D.L., Markusen, J.R., Maskus, K.E. (2001). "Estimating the knowledge capital model of the multinational enterprises". *American Economic Review* 91: 693-708.

Chaney, T. (2011). "The Network Structure of International Trade," *NBER Working Papers*, 16753, National Bureau of Economic Research.

Disdier, A.C., Head, K. (2008). "The puzzling persistence of the distance effect on bilateral trade. Review of Economics and Statistics 90:37-48.

Djankov, S., La Porta, R., Lopez-de-Silanes, F., Shleifer, A. (2002). The regulation of entry. *Quarterly Journal of Economics* 117:1-37.

Drogué, S., Bartova, L. (2007). "A critical survey on tariffs and trade available for the analysis of EU agriculture agreements". TradeAg Working Paper 07/6.

Dunning, J.H. (1977). "Trade, Location of Economic Activity and MNE: A Search for an Eclectic Approach" In: *International Allocation of Economic Activity*, eds. Hohlin, Hesselborn and Wijkman. McMillan.

Francois, J. (2010). "Who trades with whom". Paper presented at the 12th European Trade Study Group, Lausanne 9-11 September.

Greenaway, J. Kneller, J. (2007). "Firm heterogeneity, exporting and foreign direct investment". *The Economic Journal* 117:134-161.

Head, K., Ries, J. (2008). "FDI as an outcome of the market for corporate control: Theory and evidence". (2008). *Journal International Economics* 74: 2-20.

Helpman, E., M.J. Melitz, S.R. Yeaple. (2004). "Export versus FDI with heterogeneous firms," *American Economic Review*, 94(1), 300-16.

Helpman, E., Melitz, M.J., Rubistein, Y. (2008). "Estimating trade flows: trading partners and trading volumes" *The Quarterly Journal of Economics* 73:441-487.

Herger N., C. Kostoggiannis, S. McCorriston. (2008). “Cross-border acquisitions in the global food sector”. *European Review of Agricultural Economics* 35(4), 563-587.

Isaksson, A. (2009). “The UNIDO world productivity database: an overview”. *International Productivity Monitor* 18, 38–50.

Johnson, D. K. N. (2002), The OECD Technology Concordance (OTC): Patents by industry of manufacture and sector of use. OECD Science, Technology and Industry Working Papers 2002/5, OECD, Directorate for Science, Technology and Industry.

Laursen K. (1998). “Revealed comparative advantage and the alternatives as measures of international specialization”. Danish Research Unit for Industrial Dynamics Working Paper n. 98-30.

Markusen, J.R., Maskus, K. (2002). “Discriminate among alternative theories of the multinational enterprise. *Review of International Economics* 10:694-707.

Melitz, M.J. (2003). “The impact of trade on intra-industry reallocation and aggregate industry productivity”. *Econometrica* 71(6), 1695-1725.

Michaely, M. (1967). “Concentration in international trade, contributions to economic analysis”. Amsterdam, North-Holland Publishing Company.

Oldenski, L. (2010). Export versus FDI: a task based approach. Mimeo.

Slangen, A.H.L., Beugelsdijk, S. (2010). The impact of institutional hazards on foreign multinational activity: A contingency perspective. *Journal of International Business Studies*, 41, 980–995.

Slangen, A.H.L., Beugelsdijk, S., Hennart, J.M.A. (2011). The impact of cultural distance on bilateral arm's length exports: an international business perspective. *Management International Review*, 51(6), 875-896.

Tomiura, E. (2007). “Foreign outsourcing, exporting, and FDI: A productivity comparison at the firm level.” *Journal of International Economics*, Elsevier, vol. 72(1), pages 113-127.

UNCTAD (1999). World Investment Report 1999. United Nations Conference on Trade and Development, Geneva.

UNCTAD (2010). World Investment Report 2010. United Nations Conference on Trade and Development, Geneva.

Wang, C., Wei, Y., Liu, X. (2010). “Determinants of bilateral trade flows in OECD countries: evidence from gravity panel data models”. *The World Economy* 33:894-915.

Yeaple, S. (2003). “The role of skill endowments in the structure of US outward foreign investments. *Review of Economics and Statistics* 85:726-734.

Table 1 - Variables description and sources

Description and sources of all the variables used in the empirical analysis..

Definition	Description and Source
Exports	Value of exports from country *i* to country *j* in sector *h*. *Source*: UN Comtrade
FDIs	Value of mergers and acquisitions from country *i* to country *j* in sector *h*. *Source*: SDC Platinum
Num. of large firms	Number of firms in country *i* in the first decile of the world distribution of firm sales in a given sector *h*. *Source:* Worldscope Database
Num. of large firms in other countries	Number of large firms in a sector *h*, in all countries except for country *i*. *Source:* Worldscope Database
Num. of large firms in other sectors	Number of large firms in a country *i*, in all sectors except for sector *h*. *Source:* Worldscope Database
TFP (log)	Log of average level of total factor productivity in sector *h* in country *i*. *Source*: UNIDO (Indstat4, 2008 version)
Wage (log)	Log of average wages in sector *h* in country *i*. *Source*: UNIDO (Indstat4, 2008 version)
Capital intensity	Ratio between capital and number of employees in sector *h* in country *i*. *Source*: UNIDO (Indstat4, 2008 version)
Patents	Number of patents produced in a country *i* and in a given sector *h* and granted by the US Patent Office. *Source*: NBER
Distance (log)	Log of average distance between countries *i* and *j* calculated through the great circle formula that uses latitudes and longitudes of the most important cities (in terms of population). *Source*: CEPII http://www.cepii.fr/anglaisgraph/bdd/distances.htm
Islands	Number of countries that are islands in the pair of countries *i* and *j*. *Source*: CEPII http://www.cepii.fr/anglaisgraph/bdd/distances.htm
Landlocked	Number of countries that are landlocked in the pair of countries *i* and *j*. *Source*: CEPII http://www.cepii.fr/anglaisgraph/bdd/distances.htm
Common legal system	Dummy variable equal to 1 if country *i* and *j* share the same legal system. *Source*: CEPII http://www.cepii.fr/anglaisgraph/bdd/distances.htm
Common language	Dummy variable equal to 1 if country *i* and *j* share the same language. *Source*: CEPII http://www.cepii.fr/anglaisgraph/bdd/distances.htm
Common religion	Dummy variable equal to 1 if country *i* and *j* share the same religion. *Source*: CEPII http://www.cepii.fr/anglaisgraph/bdd/distances.htm
Contiguity	Dummy variable equal to 1 if country *i* and *j* share common borders. *Source*: CEPII http://www.cepii.fr/anglaisgraph/bdd/distances.htm
Colonial ties	Dummy variable equal to 1 if country *i* and *j* have ever been in colonial relationship. *Source*: CEPII http://www.cepii.fr/anglaisgraph/bdd/distances.htm
Tariffs	Tariffs applied from country *j* to country *i* in sector *h*. *Source*: TRAINS
Common partners in trade	Number of partners in trade common to country *i* and *j* in sector *h*. *Source*: UN Comtrade
Common partners in FDIs	Number of partners in FDIs common to country *i* and *j* in sector *h*. *Source*: SDC Platinum
Share greenfield	Number of greenfield investment realized in a given destination country *j* relative to total number of greenfield investment in the world. *Source*: UNCTAD

Table 2 – Summary statistics

Variable	Mean	Median	Standard deviation	Min	Max
Index_Norm	0.051	0.008	0.188	-0.884	7.058
Num. of large firms	3	1	3	1	55
Num. of large firms in other countries	44	38	32	1	167
Num. of large firms in other sectors	120	25	232	0	763
TFP (log)	5.073	5.174	0.705	1.614	7.785
Wage (log)	10.054	10.294	1.230	6.279	12.559
Capital intensity	1.681	1.664	0.178	1.309	2.468
Patents	17	0	89	0	1,465
Distance (log)	8.826	9.052	0.742	5.371	9.892
Islands	0.417	0	0.570	0	2
Landlocked	0.164	0	0.383	0	2
Common legal system	0.277	0	0.447	0	1
Common language	0.108	0	0.310	0	1
Common religion	0.197	0.040	0.291	0	0.988
Contiguity	0.021	0	0.143	0	1
Colonial ties	0.039	0	0.194	0	1
Tariffs	11.734	9.295	10.770	0	58.235
Common partners in trade	58	58	37	0	117
Common partners in FDIs	0	0	1	0	30
Share greenfield	1.124	0.464	2.376	0.000	17.608

Notes: Variables description and sources are provided in Table 1. Summary statistics are computed after excluding observations in the 1st and the 99th percentile of the distribution of the dependent variable. Summary statistics are calculated on 67,911 observations for all variables.

Table 3 – Correlation matrix

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
(1) *Index_Norm*	1																			
(2) *Num. of large firms*	-0.004	1																		
(3) *Num. of large firms in other countries*	0.076	0.032	1																	
(4) *Num. of large firms in other sectors*	-0.054	0.792	-0.189	1																
(5) *TFP (log)*	0.073	0.340	-0.113	0.359	1															
(6) *Wage (log)*	0.062	0.352	0.082	0.292	0.678	1														
(7) *Capital intensity*	-0.055	-0.157	-0.103	-0.137	0.384	0.365	1													
(8) *Patents*	-0.025	0.422	-0.043	0.459	0.117	0.146	0.108	1												
(9) *Distance (log)*	0.014	0.065	-0.015	0.088	-0.065	-0.098	-0.098	0.029	1											
(10) *Islands*	-0.042	0.153	-0.054	0.163	0.181	0.146	0.094	-0.001	0.067	1										
(11) *Landlocked*	0.013	-0.029	0.027	-0.039	0.029	0.075	0.070	-0.014	-0.105	-0.137	1									
(12) *Common legal system*	0.011	-0.087	0.006	-0.082	-0.110	-0.130	-0.084	-0.010	-0.043	0.021	-0.071	1								
(13) *Common language*	-0.011	0.022	-0.018	0.050	-0.014	-0.056	-0.071	0.059	-0.102	0.113	-0.024	0.409	1							
(14) *Common religion*	0.030	-0.116	0.004	-0.112	-0.045	-0.059	-0.013	-0.029	-0.064	-0.073	0.041	0.279	0.103	1						
(15) Contiguity	-0.006	-0.039	-0.002	-0.032	-0.033	-0.047	-0.015	-0.006	-0.394	-0.089	0.054	0.123	0.146	0.122	1					
(16) *Colonial ties*	-0.018	0.036	-0.030	0.021	0.047	0.070	-0.025	0.005	-0.031	0.212	-0.047	0.236	0.323	-0.028	0.015	1				
(17) *Tariffs*	-0.003	0.007	-0.124	0.037	0.028	0.031	-0.018	0.011	0.018	-0.062	0.016	0.031	0.021	-0.144	-0.038	-0.005	1			
(18) *Common partners in trade*	0.005	0.013	0.091	-0.008	-0.068	0.021	-0.122	0.019	-0.074	-0.022	-0.109	-0.108	-0.071	-0.010	0.055	0.022	-0.230	1		
(19) *Common partners in FDIs*	-0.028	0.218	0.148	0.150	0.091	0.146	-0.019	0.117	-0.036	0.024	0.077	-0.054	0.000	0.140	0.087	0.015	-0.171	0.334	1	
(20) *Share greenfield*	-0.021	-0.025	-0.015	-0.022	-0.009	-0.017	0.012	-0.011	-0.031	-0.092	-0.113	-0.095	-0.004	-0.127	0.043	0.007	0.048	0.342	0.096	1

Notes. Variable definitions and sources are provided in Table 1. Correlations are computed after excluding observations in the 1st and the 99th percentile of the distribution of the dependent variable. Correlations are calculated on 67,911 observations for all variables.

Table 4 – Baseline regression and sample splits

Variables description and sources are provided in Table 1. The dependent variable is *Index_Norm*. All columns report two-stage least-squares estimates, instrumenting *Num. of large firms* (*log*) through *Num. of large firms in other countries* and *Num. of large firms in other sectors.* Column 1 reports estimates on the whole sample. Columns 2 and 3 report estimates on the subsamples of sectors with wages above and below the median level, respectively. Columns 4 and 5 report estimates on the subsamples of sectors with capital intensity above and below the median level, respectively. Coefficients of the following variables are multiplied by 1,000: *Patents*, *Common partners in trade* and *Common partners in FDIs.* All estimates include unreported domestic country, foreign country and sector-specific fixed effects. Standard errors, robust to heteroscedasticity, are reported in parenthesis. ***, ** and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

	(1)	(2)	(3)	(4)	(5)
		Wages		Capital intensity	
	All sample	Above the median	Below the median	Above the median	Below the median
Num. of large firms (log)	-0.0116***	-0.0213***	0.0032*	-0.0472***	0.0147***
	(0.0025)	(0.0047)	(0.0019)	(0.0054)	(0.0031)
TFP (log)	0.0595***	0.0637***	0.0464***	0.0170***	0.1085***
	(0.0038)	(0.0050)	(0.0039)	(0.0031)	(0.0078)
Wage (log)	-0.0067*	0.0089**	-0.0180***	0.0163***	-0.0184***
	(0.0036)	(0.0036)	(0.0054)	(0.0031)	(0.0060)
Capital intensity	-0.1170***	-0.2130***	-0.0625***	-0.0648***	-0.4292***
	(0.0062)	(0.0121)	(0.0066)	(0.0068)	(0.0245)
Patents	0.0159***	0.0274***	-0.0058*	0.1448***	0.0039
	(0.0032)	(0.0056)	(0.0032)	(0.0265)	(0.0035)
Distance	0.0076***	0.0089***	0.0052***	0.0011	0.0136***
	(0.0013)	(0.0019)	(0.0015)	(0.0012)	(0.0021)
Islands	0.0290	0.0259	0.0287**	-0.0047	0.0579***
	(0.0181)	(0.0304)	(0.0137)	(0.0202)	(0.0172)
Landlocked	0.0250*	0.0527**	-0.0134	-0.0027	0.0515
	(0.0147)	(0.0249)	(0.0116)	(0.0206)	(0.0333)
Common legal system	0.0011	-0.0023	0.0054*	-0.0044**	0.0053
	(0.0021)	(0.0029)	(0.0028)	(0.0021)	(0.0035)
Common language	-0.0109***	-0.0100**	-0.0106***	-0.0065**	-0.0135***
	(0.0031)	(0.0046)	(0.0039)	(0.0032)	(0.0050)
Common religion same	-0.0195***	-0.0175**	-0.0167**	0.0055	-0.0369***
	(0.0060)	(0.0085)	(0.0073)	(0.0055)	(0.0099)
Contiguity	-0.0074	-0.0065	-0.0081	0.0015	-0.0154**
	(0.0046)	(0.0080)	(0.0053)	(0.0051)	(0.0076)
Colonial ties	0.0008	-0.0050	0.0056	0.0017	-0.0003
	(0.0033)	(0.0053)	(0.0037)	(0.0046)	(0.0049)
Tariffs	-0.0765***	-0.0607***	-0.0571***	-0.0387***	-0.1274***
	(0.0135)	(0.0172)	(0.0193)	(0.0120)	(0.0278)
Common partners in trade	0.9306***	0.9697***	0.6643***	0.5001***	1.2635***
	(0.0662)	(0.1005)	(0.0772)	(0.0585)	(0.1144)
Common partners in FDIs	-2.5984***	-2.4003***	-3.2399***	-2.3795***	-2.9619***
	(0.4165)	(0.6183)	(0.4732)	(0.6650)	(0.5665)
Observations	67,911	35,528	32,383	32,473	35,438
Adjusted R^2	0.141	0.174	0.247	0.113	0.185

Table 5 – Groups of countries

Variables description and sources are provided in Table 1. The dependent variable is *Index_Norm*. All columns report two-stage least-squares estimates, instrumenting *Num. of large firms* (*log*) through *Num. of large firms in other countries* and *Num. of large firms in other sectors*. Column 1 reports estimates on the subsample of origin countries belonging to the group of G-10 countries. Column 2 reports estimates on the subsamples of origin countries belonging to the group of OECD countries. Column 3 reports estimates on the subsample of origin and destination countries belonging to the group of OECD countries. Column 4 reports estimates on the subsample of destination countries belonging to the group of developing countries. Columns 5 and 6 report estimates on the subsamples of destination countries belonging to the group of countries with the share of greenfield investment above and below the median level, respectively. Coefficients of the following variables are multiplied by 1,000: *Patents*, *Common partners in trade* and *Common partners in FDIs*. All estimates include unreported domestic country, foreign country and sector-specific fixed effects. Standard errors, robust to heterosckedasticity, are reported in parenthesis. ***, ** and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
					Share of greenfield investment	
	G-10 (origin)	OECD (destination)	OECD (origin and destination)	Developing (destination)	Above the median	Below the median
Num. of large firms (log)	-0.0411***	-0.0223***	-0.0136***	-0.0100***	-0.0147***	-0.0081**
	(0.0033)	(0.0028)	(0.0041)	(0.0036)	(0.0034)	(0.0036)
TFP (log)	0.0839***	0.0774***	0.0529***	0.0703***	0.0516***	0.0706***
	(0.0078)	(0.0051)	(0.0095)	(0.0059)	(0.0047)	(0.0062)
Wage (log)	0.0641***	0.0439***	0.0379***	-0.0038	-0.0051	-0.0099*
	(0.0060)	(0.0040)	(0.0064)	(0.0054)	(0.0042)	(0.0059)
Capital intensity	-0.1533***	-0.0851***	-0.0984***	-0.1077***	-0.1363***	-0.1271***
	(0.0123)	(0.0086)	(0.0144)	(0.0091)	(0.0087)	(0.0103)
Patents	0.0144***	0.0115***	0.0049	0.0103**	0.0242***	0.0055
	(0.0050)	(0.0033)	(0.0055)	(0.0041)	(0.0044)	(0.0047)
Distance	0.0038**	0.0060***	0.0026	0.0095***	0.0071***	0.0082***
	(0.0015)	(0.0014)	(0.0028)	(0.0023)	(0.0016)	(0.0021)
Islands	-0.0031	0.0019	0.0054	-0.0348***	-0.0013	-0.0511***
	(0.0135)	(0.0099)	(0.0114)	(0.0107)	(0.0093)	(0.0106)
Landlocked	-0.0076	-0.0021	-0.1231***	0.0241	0.1889***	0.0169**
	(0.0172)	(0.0112)	(0.0303)	(0.0147)	(0.0177)	(0.0084)
Common legal system	-0.0049*	-0.0030	-0.0095**	0.0044	0.0012	0.0004
	(0.0029)	(0.0021)	(0.0040)	(0.0034)	(0.0027)	(0.0033)
Common language	0.0001	-0.0023	0.0097	-0.0189***	-0.0124***	-0.0091**
	(0.0039)	(0.0030)	(0.0064)	(0.0056)	(0.0043)	(0.0045)
Common religion same	-0.0055	-0.0052	0.0084	-0.0179**	-0.0151*	-0.0259***
	(0.0076)	(0.0049)	(0.0090)	(0.0072)	(0.0087)	(0.0082)
Contiguity	-0.0014	-0.0011	-0.0038	-0.0178**	-0.0055	-0.0125
	(0.0083)	(0.0050)	(0.0062)	(0.0077)	(0.0060)	(0.0080)
Colonial ties	-0.0008	0.0011	0.0010	0.0015	-0.0017	0.0087*
	(0.0045)	(0.0035)	(0.0073)	(0.0050)	(0.0042)	(0.0051)
Tariffs	-0.0866***	-0.0768***	0.0158	-0.1171***	-0.0561***	-0.0823***
	(0.0180)	(0.0126)	(0.0374)	(0.0184)	(0.0161)	(0.0210)
Common partners in trade	0.3878***	0.3956***	0.2700**	0.4823***	1.3605***	0.6903***
	(0.0712)	(0.0553)	(0.1321)	(0.0835)	(0.1018)	(0.0914)
Common partners in FDIs	-3.1769***	-3.2812***	-1.8330***	-13.8987***	-1.9026***	-2.7921**
	(0.4733)	(0.4215)	(0.5142)	(4.4635)	(0.4464)	(1.2048)
Observations	23,430	41,631	10,453	37,843	34,687	33,060
Adjusted R^2	0.175	0.173	0.175	0.143	0.162	0.131

Table 6 – Heckman models

Variables description and sources are provided in Table 1. Column 1 reports estimates of the first probit model. The dependent variable in Column 1 takes the value of one if sector *h* of country *i* exports and/or invests in country *j*, and the value of zero if both exports and FDIs are zero for a sector *h* of country *i*. Column 2 reports two-stage least-squares estimates on the subsample including all cases in which sector *h* of country *i* exports and/or invests in country *j*, instrumenting *Num. of large firms* (*log*) through *Num. of large firms in other countries* and *Num. of large firms in other sectors*. Column 3 reports estimates of the second probit model. The dependent variable in Column 3 takes the value of one if sector *h* of country *i* exports and invests in country *j*, and the value of zero otherwise. Column 4 reports two-stage least-squares estimates on the subsample including all cases in which sector *h* of country *i* exports and invests in country *j*, instrumenting *Num. of large firms (log)* through *Num. of large firms in other countries* and *Num. of large firms in other sectors*. Coefficients of the following variables are multiplied by 1,000: *Patents, Common partners in trade and Common partners in FDIs.* All estimates include domestic country, foreign country and sector-specific fixed effects (not reported). In columns 1 and 3 standard errors, reported in parenthesis, are clustered by country pairs. In columns 2 and 4 standard errors, reported in parenthesis, are bootstrapped with 100 replications. ***, ** and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

	(1)	(2)	(3)	(4)
	Heckman model 1		Heckman model 2	
	Probit	Two-stage least-squares	Probit	Two-stage least-squares
Number of large firms (log)	0.2401***	-0.0099***	0.1917***	-0.0202***
	(0.0397)	(0.0027)	(0.0284)	(0.0052)
TFP (log)	0.1390***	0.0717***	-0.0396	0.0412***
	(0.0387)	(0.0045)	(0.0584)	(0.0104)
Wage (log)	0.0941**	-0.0129***	0.1091*	0.0222**
	(0.0415)	(0.0042)	(0.0612)	(0.0095)
Capital intensity	-0.8710***	-0.1263***	-0.0583	-0.1110***
	(0.1137)	(0.0069)	(0.1595)	(0.0278)
Patents	0.5265	0.0161***	-0.0598	0.0023
	(0.3834)	(0.0030)	(0.0923)	(0.0088)
Distance	-0.8398***	0.0101***	-0.4544***	-0.0013
	(0.0607)	(0.0015)	(0.0424)	(0.0039)
Islands	0.0299	-0.0010	-0.4889	0.0052
	(0.2979)	(0.0100)	(0.3021)	(0.0171)
Landlocked	-1.0099***	0.0938***	-1.1798***	0.0023
	(0.1953)	(0.0116)	(0.4409)	(0.0308)
Common legal system	0.0479	0.0021	-0.0099	0.0027
	(0.0656)	(0.0026)	(0.0891)	(0.0057)
Common language	0.6494***	-0.0145***	0.4201***	-0.0022
	(0.1002)	(0.0038)	(0.1194)	(0.0076)
Common religion same	0.2061**	-0.0261***	0.4915*	0.0266
	(0.0976)	(0.0078)	(0.2626)	(0.0298)
Contiguity	-0.5860***		0.1741	
	(0.2070)		(0.1318)	
Colonial ties	0.6572**		0.2052*	
	(0.2688)		(0.1163)	
Tariffs	-0.4899***	-0.0918***	0.5827**	-0.0062
	(0.1837)	(0.0131)	(0.2720)	(0.0368)
Common partners in trade	16.4026***	1.0254***	11.8321***	-0.1651
	(1.4297)	(0.0741)	(1.4636)	(0.2006)
Common partners in FDIs	159.8385**	-2.3915***	136.1897***	-1.6001***
	(66.7379)	(0.3898)	(12.9980)	(0.5601)
Mills ratio		0.0517***		-0.0051
		(0.0123)		(0.0103)
Observations	63,176	60,298	48,167	3,755
Adjusted R^2		0.164		0.259